AN ATTEMPT AT STANDARDIZATION OF JUNIOR SECONDARY THREE ALGEBRA TESTS USING THE CONTEMPORARY MATHEMATICS CURRICULUM IN KANO STATE OF NIGERIA.

BY

IDRISU UMARU JAHUN

A THESIS SUBMITTED IN PARTIAL FULFIILMENT OF THE REQUIREMENT FOR MASTERS'S DEGREE

IN

CURRICULUM AND INSTRUCTION
(MATHEMATICS)

DEPARTMENT OF EDUCATION,
AHMADU BELLO UNIVERSITY,
ZARIA.

1985.

DECLARATION.

I hereby certify that this thesis has been written by me and that it is a record of my own research work. It has not been presented in any previous application for a higher degree. All quotations are indicated by quotation marks or indentation and the sources of information are specifically acknowledged by means of references.

CERTIFICATION

This thesis entitled "AN ATTEMPT AT STANDARDIZATION OF JUNIOR SECONDARY THREE ALGEBRA TESTS USING THE CONTEMPORARY MATHEMATICS CURRICULUM IN KANO STATE OF NIGERIA" by Idrisu Umaru Jahun meets the regulations governing the award of the degree of M.Ed (Curriculum and Instruction in Mathematics) of Ahmadu Bello University, and is approved for its contribution to knowledge and literary presentation.

Dr. A.A. Sambo
Chairman, Supervisory Committee — Date: 13/8/86

Dr. L. Shirley
Member, Supervisory Committee — Date: _____

Mr. Y. Pathan
Member, Supervisory Committee — Date: 13/8/86

Dr. A.I. Ikeotuonye
Member, Supervisory Committee — Date: 13/8/86

Dr. A.A. Sambo
Head of Department — Date: 13/8/86

Prof. G.E. Osuide
Dean, Postgraduate School — Date: 29/ix/86

DEDICATION.

This work is dedicated to my late father, Alhaji Umaru Mohammed Jahun who inspired me to read Mathematics at University level.

DEPARTMENT OF EDUCATION
AHMADU BELLO UNIVERSITY
ZARIA.

COPYRIGHT

Name of Candidate: IDRISU UMARU JAHUN

Tittle of Thesis: An Attempt at Standardization of Junior Secondary Three Algebra Tests Using the Contemporary Mathematics Curriculum in Kano State of Nigeria.

Degree for which report was submitted: Master's in Curriculum & Instruction in Mathematics.

Year this degree was granted:

Permission is hereby given to the Ahmadu Bello University, Zaria to produce single copies of this report for Scholarly or Scientific research purpose only.

The author reserves other publication rights, and neither this report nor extensive extracts from it may be printed or reproduced without the author's written permission.

IDRISU UMARU JAHUN.

ACKNOWLEDGEMENT

It is not possible in a study of this magnitude to pay adequate tribute to all those who have contributed directly or indirectly towards its success. However, I would like to express my appreciation and gratitude to Dr. A.A. Sambo, Head, Department of Education and Dean of the Faculty of Education of Ahmadu Bello University, Zaria who encouraged me to embark in a study of this nature. Dr. Sambo, who served as my supervisor, persistantly encouraged and advised me on the study. I am particularly grateful to him for finding time to read the drafts of this study- inspite of his administrative workload-and for his constructive criticisms and suggestions.

My thanks are due to Mr. Y. Pathan of the Faculty of Education, Ahmadu Bello University, Zaria for his painstaking guidance in the orientation and statistical design of this study. I am also grateful to Professor S.C. Aleyideino of the Institute of Education, Ahmadu Bello University, Zaria, who continuously urged me to work harder on the study and for reading part of the manuscripts. My thanks also go to the entire staff members of the Computer Centre, Ahmadu Bello University, Zaria for their help in processing the data of this study.

I cannot hold any of the above persons responsible for the final version of this thesis.

Special thanks are due to Kano State Government for its moral and financial support which made this work possible. I am particularly appreciative of the aids given to me by the Chief Inspector of Education, Kano Educational Resource Centre, Alhaji Lawal Na-Allah, in conducting the workshop and for administering and scoring the tests. I am also thankful to the students, teachers and various school authorities for the maximum cooperation I enjoyed from them. Special mention should also be made here of the grant awarded to me by Department of Education and Post-Graduate School both of Ahmadu Bello University, Zaria which considerably facilitated the write-up of this thesis.

Mr. Noel M. Agbayi of Department of Education, Ahmadu Bello University, Zaria, accepted the task of typing and retyping the drafts in spite of my illegible handwriting. His untiring efforts made it possible to convert, neatly, the manuscript of this study into typed bundle of papers. His assistance is highly appreciated.

Finally, I must express my gratitude to my wives, Talatu and Sa'adatu, who wonderfully endured my absence and for taking care of our children at the time I was deeply involved in the study.

ABSTRACT

The main objective of this study was to construct and standardize . tests on Secondary Three Algebra Course, The study made use of the contemporary-Junior Secondary School Mathematics Curriculum.

Five sets of parallel tests were constructed and administered to 349 Form Three Students and 659 begining Form Four Students in Kaduna and Kano States, respectively, for the first and second try-out. On the basis of item analysis, two sets of parallel tests were finally developed and administered to 3081 Form Three Students in 14 Schools randomly selected in Kano State of Nigeria. An evaluation model was identified for assessing achievement in secondary algebra which was used to construct the tests.

From the study data Three different norms have been established. These are; norms for the total sample, sex norma and Institution type norms for each test set. Altogether ten different norms have been identified.

Each of the ten norms identified has been presented and discussed by the use of graphs, percentile ranks and normalized standard scores.

The study showed that the mean performance of Girls is higher than the mean performance of Boys in each test set. Also the mean performance of Private Institutions is very much higher than the mean performance of Government Institutions in each test set.

TABLES OF CONTENTS

		Page
Title Page	::::	i
Declaration	::::	ii
Certification	::::	iii
Dedication	::::	iv
Copyright	::::	v
Acknowledgement	::::	vi
Abstract	::::	viii
Tables of Contents	::::	x
List of Tables	::::	xiii
List of Figures	::::	xvi
List of Appendices	::::	xvii
Definition of Terms	::::	xviii
Chapter 1 THE PROBLEM		
1.1 Introduction	::::	1
1.2 Purpose of the Study	::::	3
1.3 Significance of the Study	::::	5
1.4 Scope and Limitation of the Study		7
1.5 Objectives of the Study	::::	9
Chapter 2 LITERATURE REVIEW AND THEORETICAL FRAMEWORK		
2.1 Introduction	::::	11
2.2 Evaluation of Students	::::	12
2.3 Evaluation Models	::::	16
2.4 Evaluation Procedures	::::	26
2.5 Evaluation Instruments	::::	30
2.6 Types of Tests	::::	33
2.7 Standardized Tests	::::	45

Tables of Contents Contd. Page

2.8 Theoretical Aspect of Standardization Process of a Test :::: 48

(a) Planning the Test :::: :::: 50
(b) Writing the Items :::: :::: 51
(c) Pretesting Items :::: :::: 52
(d) Preparing the Final Form(s) :::: 53
(e) Psychometric Qualities :::: 55
(f) Norms :::: :::: :::: 59
(g) Test Manuals :::: :::: 62

2.9 Empirical Studies :::: :::: 62

Chapter 3 DESIGN OF THE STUDY

3.1 The Pupulation :::: :::: :::: 78

3.2 Sample :::: :::: :::: 79

i. The Samples For The First and Second Try-out :::: :::: 80

ii. The Sample For The Final Testing 82

3.3 Design Implementation :::: :::: 83

Chapter 4 DEVELOPMENT OF THE DATA COLLECTION INSTRUMENTS

4.1 Introduction :::: :::: :::: 85

4.2 Items Constructed :::: :::: 86

4.3 First Try-out :::: :::: :::: 93

4.4 The Workshop :::: :::: :::: 95

4.5 Second Try-out :::: :::: :::: 97

i. Administration :::: :::: 98
ii. The Results :::: :::: 99
iii. Statistical Analysis :::: 99

4.6 The Final Test :::: :::: :::: 105

4.7 Statistical Analysis Of Final Testing 107

Tables of Contents Contd. Page

Chapter 5 DATA PRESENTATION AND ANALYSIS

5.1 Introduction :::: :::: :::: 114

5.2 Frequency Distribution Of The Study Data :::: :::: :::: 114

5.3 Percentile Ranks Of The Test Scores :::: 129

5.4 Normalized Standard Scores Of The Distribution :::: :::: :::: 130

5.5 Reliability Estimate :::: :::: 131

5.6 Standard Error Of Measurement :::: 134

Chapter 6 DISCUSSION, MANUAL INFORMATION AND RECOMMENDATION

6.1 Discussion :::: :::: :::: 135

6.2 Manual Information :::: :::: 138

6.3 Recommendations For Further Research 147

BIBLIOGRAPHY :::: :::: :::. 150

APPENDICES :::: :::: :::: 160

LIST OF TABLES

			Page
Table	3.1	Institutions Selected for the First Try-out in Zaria Local Government Area of Kaduna State	81
Table	3.2	Institutions Selected for the Second Try-out in Kano State	81
Table	3.3	Institutions Selected for the Final Testing	83
Table	4.1	Guideline for Constructing the Test Items For Standardized Achievement Test For Junior Secondary 3 Algebra Tests	88
Table	4.2	Table of Specification For The Standardized Test for Junior Secondary 3 Algebra Tests.	91
Table	4.3	Significance of Means and Variance	101
Table	4.4	Computation of Percentile Rank	109
Table	4.5	Computation of Normalized Standard Scores	110
Table	5.1	Means and Standard Deviations Of The Various Norms and Test Set	115
Table	5.2	Significance Of The Difference Between Means	116
Table	5.3	Skewness and Kurtosis of Scores Distribution of the Various Norms and Test Set.	117
Table	5.4	Half-Length Reliability Coefficients of Each Test By Norms	131
Table	5.5	Reliability Coefficients of Each Test By Norms	132
Table	5.6	Standard Error of Measurement of Each Test By Norms	134

List of Tables Contd. Page

Table	6.1	Part of the Results of the First Try-out	
		Set A :::: :::: ::::	161
		Set B :::: :::: ::::	162
		Set C :::: :::: ::::	163
		Set D :::: :::: ::::	164
Table	6.2	Frequency Table Of Set A Total	219
Table	6.3	Frequency Table Of Set B Total	220
Table	6.4	Frequency Table Of Set A Boys	221
Table	6.5	Frequency Table Of Set B Boys	222
Table	6.6	Frequency Table Of Set A Girls	223
Table	6.7	Frequency Table Of Set B Girls	224
Table	6.8	Frequency Table Of Set A Private	225
Table	6.9	Frequency Table Of Set B Private	226
Table	6.10	Frequency Table Of Set A State Government	227
Table	6.11	Frequency Table Of Set B State Government	228
Table	6.12	Percentile Rank Set A Total	229
Table	6.13	Percentile Rank Set B Total	230
Table	6.14	Percentile Rank Set A Boys	231
Table	6.15	Percentile Rank Set B Boys	232
Table	6.16	Percentile Rank Set A Girls	233
Table	6.17	Percentile Rank Set B Girls	234
Table	6.18	Percentile Rank Set A Private	235
Table	6.19	Percentile Rank Set B Private	236

List of Tables Contd. Page

Table	6.20	Percentile Rank Set A State Government	237
Table	6.21	Percentile Rank Set B State Government	238
Table	6.22	Normalized Standard Scores Set A Total	239
Table	6.23	Normalized Standard Scores Set B Total	240
Table	6.24	Normalized Standard Scores Set A Boys	241
Table	6.25	Normalized Standard Scores Set B Boys	242
Table	6.26	Normalized Standard Scores Set A Girls	243
Table	6.27	Normalized Standard Scores Set B Girls	244
Table	6.28	Normalized Standard Scores Set A Private	245
Table	6.29	Normalized Standard Scores Set B Private	246
Table	6.30	Normalized Standard Scores Set A State Government	247
Table	6.31	Normalized Standard Scores Set B State Government	248

LIST OF FIGURES

Figure		Page
5.1	Frequency Polygon of Set A Total	119
5.2	Frequency Polygon of Set B Total	120
5.3	Frequency Polygon of Set A Boys	121
5.4	Frequency Polygon of Set B Boys	122
5.5	Frequency Polygon of Set A Girls	123
5.6	Frequency Polygon of Set B Girls	124
5.7	Frequency Polygon of Set A Private	125
5.8	Frequency Polygon of Set B Private	126
5.9	Frequency Polygon of Set A State Govt.	127
5.10	Frequency Polygon of Set B State Govt.	128

LIST OF APPENDICES

		Page
APPENDIX A	The Domain of the Study	160
APPENDIX B	Part of the Results Of The First Try-out	161
APPENDIX C	Standardized Test Research Information Sheet (Teachers & Schools)	165
APPENDIX D	The Five Sets Of Tests For Second Try-out	166
APPENDIX E	Item Analysis	191
APPENDIX F	Final Tests	192
APPENDIX G	Marking Scheme For The Final Tests	202
APPENDIX H	Frequency Tables Of Study Data	219
APPENDIX I	Percentile Rank Of Study Data	229
APPENDIX J	Normalized Standard Scores Of Study Data	239

DEFINITION OF TERMS

The terms below have the following operational definitions in this study.

Cognitive Achievement Test – Test that is used to assess the intellectual and academic progress of the students.

Norms – table of equivalents between raw scores and standard scores in percentile ranks and T-Scores forms: (Ebel 1979).

Standardized Test – a test which contain fixed set of items that are carefully and systematically prepared, with uniform directions for administration and scoring based on norms that are clearly defined.

Standardization – the developmental processes of standardized test.

Model – the design of the standardized test.

Private – Institutions in Kano State which are not directly administered by either the State or the Federal Ministry of Education.

Chapter 1.

THE PROBLEM

1.1 Introduction

History and research have shown that throughout the existence of the mankind people have been measuring the abilities of others. It has then been recognized that there are individual differences in achievements or performances. As a result the training of citizens to perform task needed for the continuous existence and development of society, based on the individual differences may, therefore be necessary. In this process, means may have to be provided for the placement of individuals to appropriate positions in the society. This is because as human societies grow in size and complexity, the recognition of individual differences becomes more and more reflected in the social structure of the societies. The importance of this can be observed from the fact that no society has left the development and allocation of manpower to chance. A variety of ways have been developed through performance and assessment in education to place people in their rightful position in the society.

It is, usually, not possible for any society to give every one a chance to try out every position for which one is a canditate, be it in educational system or any other enterprices. This is more so, as the training for most jobs and positions become more complex and expensive. Goslin (1967) observes that the complexity and expenses in training for most jobs result in:-

> ...the necessity for devising some way to predict which individuals are best suited for which jobs before providing them with the required training".

Suggestions have also been made on how to resolve the problem; Goslin (1967:3) goes further to say that;

> One alternative is to use performance in schools as an indicator for general ability, and to allocate opportunities for further training, leading to high status positions in the society, on the basis of this performance. It has been recognized, however, that teacher judgements of pupils are often subjective and based on characteristics unrelated to the person being evaluated. In addition, standards vary from school to school and from teacher to teacher within schools. The solution to these problems has been the development of standardized objective tests of ability which may be used to assess the capabilities of large numbers of students having varying educational back-grounds .

Hence the heavy reliance by many countries, especially developed nations, on the development and administration of standardized achievement tests as indicators of general abilities.

Measurements of abilities have been used most widely in United State of America. According to Goslin (1967) there is no society that one finds so much pre-occupied with accurate assessment of ability conbined with the allocation of opportunity for advancement in the society on the basis of ability as United States of America. He however, cautions that the only possible

exception may be the Soviet Union. By and large he claims that United States has developed more systematic techniques for assessing abilities than any other nation. And one of these techniques is the development and use of standardized achievement tests.

1.2 Purpose of The Study.

The Federal Republic of Nigeria on 12th March, 1979 (Federal Ministry of Education, Lagos) presented new Mathematics curricula; to serve Primary, Teacher Training, Junior Secondary and Senior Secondary Institutions in the country in accordance with the new Policy on Education. These new curricula in Mathematics were to be operational as from September, 1982, the beginning of 1982/83 academic year; Jahun (1982).

According to the new Policy the Junior and Senior Secondary Schools are to be comprehensive. At the end of the Junior Secondary School Education only those students who are able and are willing to have complete six year Secondary Education can proceed; National Policy on Education (1981; Revised). This means that various trades will be offered at higher level in Senior Secondary School. Also the New Policy advocates for the use of Continuous Assessment in our School System. One of the implication of this is that teachers are expected to use their own test items in conjunction with Published test where available to make for the standard

required across schools. One of the purposes of this study, therefore, was to draw up test items based on the algebra section of Junior Secondary 3 Mathematics curriculum; administer them to a sample of Secondary 3 students who completed the programme in some randomly selected schools in Kano State.

The aim of the study was to standardize the test items for use in Kano State for the purpose of implementing Continuous Assessment Programme. It could also be extended to other States of Nigeria for trial purposes and possible standardization.

Wilson (1971) observes that:-

> ...the first concern of evaluation in Mathematics learning has been, and will still continue to be, cognitive outcomes or achievements .

That is to say, in any assessment of Mathematics learning the instructor is mostly concerned with the students cognition; computational skills, comprehension, application and ability to Analyse problems in Mathematics. It should, however, be noted that there are other areas in the affective outcomes that could be tested in Mathematics, but for the purpose of this study, they were considered secondary and were therefore, not tested. It was then the purpose of this study to construct and standardize test in algebra for Secondary Three Curriculum based on student's cognitive achievement.

The test is also meant to be used for the purpose of summative evaluation of the students achievement in Secondary 3 algebra and the suitability of the curriculum to the need and the ability of the students. It was hoped to be useful for decision making by Mathematics teachers, curriculum designers, textbooks writers, test developers and education administrators for guiding and counselling students.

1.3 Significance of The Study.

Throughout his reading and discussion, with some reputable educationists and education evaluators in Nigeria; the researcher has been made to understand that there have been very few attempts to under-take a study on standardized achievement test in Nigeria. The few attempts were mostly done by WAEC for the purpose of Certification. Of recent, STAN has started organizing standardization of science achievement of students at Primary, Junior and Senior Secondary Schools across the country. They are making this effort with the aid of Ibadan Institute of International Educational Evaluation and the University of Ibadan.

The researcher feels that there is the need for more of such studies in specific fields of education in Nigeria considering the rate of political, economic, technological and academic advancement in the country. We have noted eailier in this chapter that:

no society has left the development
and allocation of manpower to chance .

And that one very significant way of giving fair chances for position to individual in the society is through standardized achievement tests. In schools, standardized achievement tests could be very significant instruments for guiding and counselling students. This is more obvious with the current education policy in the country and the variation that may exist in the implementation of Continuous Assessment strategy across schools and across states in the Federation. Thus the need for more and more study in standardized achievement tests in all subject areas.

It is the researcher's convinction that since there is a break between the Junior and Senior Secondary School - which is of course, distinct as stated in the National Policy on Education, (Revised, 1981), - there is then the need for assessing the students achievement using standard measure across schools at this level. This will enable the authorities concerned to make decision for proper placement of Students to the type of Senior Secondary School that suit the individual needs.

It is quite logical, especially in areas like Mathematics to aim at a standard of measure so that individual student will be properly placed in the right type of Senior Secondary School that will suit his

ability. It is quite clear from the National Policy on Education (1981) that after the 3 year Junior Secondary School the students will be placed for Technical, Commercial, Academic Senior Secondary Schools and other fields of education. It is therefore very important to measure the standard of mathematics required by each of these areas. This is more so due to the fact that some of these areas need more mathematics than others.

The need for standard measure of mathematics achievement at this level is also made more significant by publishing a newer curriculum after the Onitsha Conference of 1981 for further Mathematics. The Curriculum is meant to be used in Senior Secondary School by the mathematically inclined. Thus, calling for pre-assessment using standard measuring Instruments for the purposes of counselling the students on which of the curriculum they should take.

1.4 Scope and Limitations of the Study.

Standardized achievement test in Mathematics comprises a vast area of research from Pre-Primary to Primary to Teacher Training Colleges, Junior and Senior Secondary Schools. And at each class of these systems standardized test of achievement is very important. Not only that, for proper assessment of students, standardized achievement tests must be prepared at each class but also all aspects of mathematics taught at

particular class should be evaluated; Jahun (1982). This, unfortunately has not been the case in this research. The study focuses only on the algebraic content of the Junior Secondary 3 curriculum. This is because the time available and the financial resources of the reseacher are very limited. But for proper assessment of both the curriculum and the students, all the aspects of the curriculum must be examined.

The part of the curriculum under study is stated in Appendix A.

The study would also be limited to only those institutions that the researcher found to have continuous mathematics teachers from form I to III in Kano State. This decision was taken as a result of the responses the researcher got from his attempt to undertake a pilot study of the research in Kaduna State. The results indicated that most of the schools tested did poorly because the students were not taught for one or two years out of their 3 years stay in school. This was discovered as a result of informal interview with the teachers and students by the researcher. (See chapter 3 and 4 for more detail).

The results of this study would only apply directly to students in Kano State. Any generalization attempted for other students in Nigeria would only be considered as a rough extrapolations of the data obtained.

1.5 Objectives of The Study.

From the algebra curriculum of Secondary 3 (Appendix A) the researcher has categorized the content into 3 sub-topics. These are:-

1. Factorization,
2. Equations, and
3. Variations.

The main theme of this study was to construct and standardize some test items from the contemporary Secondary 3 Algebra content based on the above category as they related to the objectives of the curriculum. Specific objectives of this study were:-

i. To construct tests based on the contemporary Secondary Three Algebra content.

ii. To try out the tests with a reasonable sample of the pupulation under consideration for the purposes of pilot study and developing the psychometic qualities of the tests.

iii. To revise and edit the test items based on the data in (ii).

iv. To administer the tests to a larger sample as the representative of the pupulation under study and find the standard scores.

v. To provide standard scores based on class norms, sex, and institution type

vi. To provide information which could be used to prepare a manual for the use of the tests and interpretation of the tests results.

Chapter 2.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK.

2.1 Introduction

The purpose of this chapter is to review the literature and theoretical framework related to:-

(a) evaluation in general;

(b) evaluation models of Bloom, Avital and Shettleworth and others so that evaluation model for this study will be developed;

(c) evaluation of mathematics achievement and standardization; and

(d) studies conducted within and outside Nigeria on standardization of students' achievement in mathematics and other fields of knowledge.

More specifically standardized achievement test in mathematics learning for the purpose of summative evaluation will be reviewed. The researcher has observed through his reading and discussion with educationists and evaluators, that so far in Nigeria there have been very few attempts to standardize achievement tests within our school system.

The motivating factor of this research, therefore, has been my supervisor's as well as my convintion that, there is the need for a study in this area in Nigeria

as indicated earlier. Thus the writings in this chapter form part of the justification of this research.

2.2 Evaluation of Students.

In the course of school activities day by day Adams (1966) observes, teachers, principals and other school personnel make various decisions about students and encourage them to make decisions for themselves. Decisions can be made when based on a good deal of information. Schools must have accumulated and valuable data about each student in the various school activities. This is because the data that can be useful in placing students into a Vocational Secondary School may not be the same with the data that can be most helpful in deciding which students should go into a science secondary school. We therefore want data that are relevant for each decision.

Decisions involve prediction. For teachers to predict what a student can do, they need to continually evaluate his progress. Evaluation of students progress involves the use of measurement and non-measurement data plus the value judgement that the teacher can use to make sound decisions.

Evaluation from the point of view of classroom instruction is seen by Gronlund (1981; 6) as a :-

> Systematic process of determining the extent to which instructional objectives are achieved by pupils .

This definition assumes that evaluation is a systematic process. It implies that uncontrolled observation of pupils is not considered. It also assumes that the instructional objectives have previously been determined. It is therefore difficult to judge clearly the nature and extend of pupil learning without determining the objectives for which we are evaluating. Evaluation is a comprehensive and inclusive assessment strategy of students' progress. It includes all types of assessment data such as: past record of achievement, range of mental ability scores, observation, pupil conference, self-rating, students' ratings and supervisors' ratings. It involves the use of quantitative and qualitative descriptions of the data and then making value judgements concerning the worth or value of the data obtained.

On the other hand when new curriculum programme has been developed, evaluation data enable us to determine the extent to which the new programme is effective in meeting the instructional objectives for which it is designed. Basically there are two types of curriculum evaluation. The first type of curriculum evaluation is refered to as formative evaluation; the second is called the summative evaluation; Scriven, (1967) and Gronlund (1981).

The main purpose of formative evaluation is to determine the degree of mastery of a particular learning task and to point the part of the task not mastered. The outcome of the formative evaluation is used to improve both the instructional methods and materials so that greater pupil learning and a more sound formulation and reformation of a programme will result; Dressell (1976) and Gronlund (1981). Formative evaluation requires a more microscopic and diagnostic analysis of the content. The information obtained from this type of evaluation is used as feedback into the learning system so that subsequent activities for the learner can be determined. Here, the analysis of students' error responses are appropriate and useful for evaluating the instructional strategies and materials; Wilson (1971).

Summative evaluation on the other hand, is directed toward a much more general assessment of the degree to which the larger outcomes have been attained over the entire course or some substantial part of it; Bloom et - al (1971). It's main function is to appraise the overall effectiveness of a curriculum programme which helps in decision making for grading students. The data obtained in this case are not intended to serve as a basis for modifying the procedures, but as a basis for selecting the most appropriate curriculum for school use, Gronlund (1981).

Evaluation serves a variety of functions both from the instructional point of view and curriculum development point of view. Adam (1967) sees the functions of an evaluation programme as three-fold. First, the administrative and supervisory functions, in which evaluation is used to determine the needs for supervisory assistance or for a revision of curricular activities and instructional materials. Data obtained in evaluation can also be used for administrative decisions concerning the placement of students to the appropriate group, class and/or institution type. It forms the basis for transcripts of students for cases of transfer and promotion. Second, the instructional functions, which is used to:- assess students' progress towards instructional goals; give evidence to both students and parents concerning such progress; ascertain group and individual retraining needs as a basis for grouping within the class, and individualized corrective instruction; and aid in developing hypotheses concerning basic causal factors behind any deficiencies. Third, Adam sees the function of evaluation for guidance purposes. Here the data obtained from evaluation are used in: helping the counsellor to assist students in making wise decision in the area of educational and vocational guidance; helping students towards improved self-appraisal and self-direction; and assisting in the identification of students whose special

abilities or disabilities require unusual modifications in the educational programme or referral to specialists.

2.3 Evaluation Models.

Essentially most classroom teachers evaluate students' progress in the cognitive area of students learning. It should, however, be noted that other areas of affective and psychonator outcomes could be evaluated. Perhaps the first person to think of and write on areas of learning that need to be evaluated in curriculum development was Bloom. Bloom (1956) in his famous Taxonomy of educational objectives categorizes educational outcomes in the cognitive domain into six broad levels in the increasing complexity. Cognition is basically referring to the students intellectual abilities in the various school subjects like Physics, Biology, Geography, Mathematics, Hausa etc.

According to Bloom's Model of Educational Evaluation the first level of achievement in the cognition is that of knowledge. Bloom uses the term knowledge to refer to the knowledge of specifics which involves terminology and specific facts. Knowledge of ways and means of dealing with specifics involving knowledge of conventions; knowledge of trends and sequences. Knowledge of classifications and categories; knowledge of criteria; and knowledge of methodology. Knowledge of the universals and abstractions in a field which deals with

the knowledge of principles and generalizations, and knowledge of theories and structures.

The second level in Blooms' Taxonomy is that of comprehension which deals with the lowest level of understanding. It refers to a kind of apprehension such that the individual knows what is being communicated and can make use of the idea being communicated without relating it to other material or seeing its fullest implications. It involves translation such as paraphrasing mathematical ideas from verbal materials into symbolic statements or vice versa. It involves interpretation of a generalization or principle by reordering or rearranging a given material. Comprehension at this level also involve extrapolation to determine implications, consequences, effects etc. which are in accordance with the conditions described in the original communication.

The third level of the Texonomy is Application which refers to the use of abstractions in particular and concrete situations. The abstractions may be in the form of rules of procedure or generalized methods. It may also be technical principles, ideas and theories which must be remembered and applied.

The fourth level of the Texonomy according to Bloom is that of Analysis which refers to the breaking down of materials in a communication into its component

parts for better understanding. The level of knowledge expected to be achieved here is that of recognizing various principles and generalization involved in the formulation of the materials; breaking down the properties of the material and identifying its various components. This may involve the ability to identify elements which have not been clearly defined in the communications and determining the relationship between the various components of a structure.

The fifth level of the Texonomy is that of Synthesis which involves assembling components together to form a New Whole. This involves one's ability to tell a personal experience effectively, skills in writing using an excellent organization of ideas and statement, ability to propose ways of testing hypothesis, ability to formulate appropriate hypothesis based upon an analysis of the factors involved and make modification in the hypothesis in the light of new factors and consideration and also the ability to make mathematical discoveries and generalizations.

The sixth and last level in Bloom's Texonomy of educational objectives is that of Evaluation which involves judgements about the value of materials and methods for given purposes. The judgements should be based on clearly defined criteria rather than personal opinion. The criteria may be internal in which a student

evaluate the accuracy of a communication from such evidence as logical accuracy and consistency or in depending one's line of arguments logically. It may be external in which a student evaluates a material with reference to selected or remembered criteria. This may involve comparision of major theories, generalization and facts about particular cultures.

Since the advent of Bloom's model of evaluating educational objectives, educationists and evaluators have been using the same or similar model to assess students progress in the various fields of education. Among these educationists and evaluators are mathematics educators and evaluators. Avital and Shettleworth (1968), for example, feel that although Bloom's texonomy is phrased in very general terms so that it is not tied to any particular subject matter, very few examples included in the items are from Mathematics. Thus Avital adopted the major categories of the Bloom Taxonomy to describe a model that explains more precisely the levels of mathematical performance.

According to Avital and Shettleworth (1972) there are three distinct levels of mathematical thinking that constitute the teaching objectives for mathematics. The first which is the most basic is recall and recognition of the material in the form in which it was taught. This almost corresponds to the category of knowledge in

Bloom's Taxonomy which is verbatim recall and recognition. The second which is next in level of complexity is that of straightforward generalization or transfer from the learnt material to material that is similar to it. The behaviour related to performance at this level is what Avital calls Algorithmic Thinking. At this level the student is exposed to well-defined procedures or algorithms for solving problems. The student is expected to generalize from the particular situation in which he learnt the algorithm to a new situation which is not usually very remote from it. At this level also the algorithmic thinking behaviour increases in complexity as the material to which the student generalizes becomes more remote from the material in the original learning situation. The algorithmic thinking corresponds to the levels of comprehension and Application in the Bloom's Taxonomy.

The third level of cognitive complexity in Avital & Shettleworth Taxonomy is what they call Open Search; the highest level of mathematical thinking. The student behaviour at this level is not confined to operations and solution of problems for which algorithms have been learnt. He can now interpret the whole or parts of a problem and see among them new relationships which are useful to the sought-for solutions. This level subsumes the Analysis and Synthesis categories of the Bloom's Taxonomy.

In Bloom's Taxonomy Evaluation is assumed to be the most complex level of performance. In Mathematical performance. Avital and Shettleworth feel that one cannot distinguish evaluation as psychologically distinct. For example judging the correctness of a proof by an internal analysis of the steps, which by its wording seems to fall in the category of evaluation, might be an integral part of the process of proof itself and so falls within the category of Analysis and Synthesis.

Thus according to Avital and Shettleworth (1972) we have three levels of thinking in Mathematics which correspond to the first five categories of educational objectives or mathematical performance.

Thinking Process	Taxonomic Level
recognition, recall	1. knowledge
Algorithmic thinking, generalization	2. comprehension 3. application
Open Search	4. Analysis 5. Synthesis.

Another model of mathematical evaluation based on Bloom Taxonomy was developed by Wilson (1971). Wilson adopted his model from the classification developed by the staff and consultants of the School Mathematics Study Group (SMSG) in its National Longitudimal Study of Mathematics Abilities (NLSMA). The SMSG model comprises of four levels of behaviour expected in the mathematics achievement of grades seven to twelve in

three main topics of mathematics at these grades levels. The model relates mathematical content with the behaviour outcomes expected in the study of mathematics; thus

	A.O. COMPUTATION	B.O. COMPREHENSION	C.O. APPLICATION	D.O. ANALYSIS
1.0 Number System				
2.0 Algebra				
3.0 Geometry				

Wilson (1971) expanded on this model to include subcategories in each of the categories of the two dimension. The model presented the main topics of mathematics at these levels as:-

1. Number System which includes items concerned with the nature and properties of whole numbers, integers, the rational numbers, the real numbers, the complex numbers, finite number system, matrices and determinant, probability and numeration system.

2. Algebra involving algebraic expressions, algebraic sentences and their solution, and relation and functions.

3. Geometry testing measurement, geometic phenomeno, formal reasoning and coordinate system and graphs.

Wilson's model of evaluation of achievement in maths contains four levels of behaviour in the cognitive domain. They include:-

(a) Computation items which are designed to test students' knowledge of specific facts and terminology and his ability to carry out algorithm.

(b) Comprehension. This level tests students' knowledge of rules, principles, generalization and mathematical structure. It also tests his ability to transform problems elements from one model to another, to follow a line of reasoning and to read and interpret a problem.

(c) Application items require the student ability to solve routine problems, to make comparisions, to analyze data and to recognize patterns, isomorphisms and symmeties.

(d) Analysis items demand the student to demonstrate his ability to solve non-routine problems, to discover relationships, to construct proofs, to criticize proof and to formulate and validate generalizations.

Wilson goes further to expand his model by including some components for <u>affective</u> outcomes in mathematics achievement. They include:-

(e) Interest and attitudes involving items on attitude, interest, motivation, anxiety and self-control.

(f) Appreciation involving items on extrinsic and intrinsic values and operational function of mathematics.

According to Wilson's description the computation level represents the least complex behaviour which is expected of a student as outcomes of instruction in mathematics. It requires simple recall and exercises of routine manupulation with no decision making or complex memory. The comprehension level is designed to be a more complex level of behaviour than computation. But Wilson cautions that the dividing line between the two categories seems to be artificial and vague in that sometimes the computation-level behaviours are incorporated within comprehension level behaviours. This level requires some abstraction and decision making and the student ability to translate problem elements from say verbal description to symbelic representation or vice-versa.

The application level behaviours are involved with a sequence of responses from the student. They are related to the course of study and deal with activities that are routine. The transfer to new situation is minimal. The analysis level behaviour is the highest of the cognitive categories according to Wilson (1971). This is the most complex behaviour and include most of the complex behaviour of analysis, synthesis and evaluation in the Bloom's (1956) Taxonomy and what Avital and Shettleworth (1972) call "Open Search". Wilson (1971) includes non-routine problem solving, discovery experiences and creative behaviours as it relates to mathematics in this level.

Wilson expands his model to include areas in the affective outcomes in mathematics achievement. The first level he considers is that of interest and attitudes. Here it is the students' interest in Mathematic, attitudes towards mathematics, motivation to study mathematics and son on, that will be tested. The last level of behaviour in Wilson's model is that of appreciation. This level is included because we most often include objectives for appreciation in mathematics. We talk of the usefulness of mathematics for social utility, for its structure, for its power and for its beauty and above all we talk of communicating mathematics to others possibly as teacher to indicate our appreciation for mathematics.

There are other several studies that use the Taxonomy (Bloom 1956) to arrive at similar model. For example the International Study of Achievement in Mathematics (ISAM) arrived at the following level of behaviour in the achievement in Mathematics:-

1. Knowledge and Information involving recall of definitions, notations and concepts.
2. Technique and Skill involving computation and manipulation of symbol.
3. Translation of data into symbols or schema or vice versa.
4. Comprehension involving capacity to analyze problems to follow reasoning.
5. Inventiveness involving reasoning creatively in Maths.

(Husen, 1967).

Wood (1968) on the other hand uses the Taxonomy and the ISAM model to arrive at his model. The model has the same content with levels 1, 2 and 5 of the ISAM Model. Wood's categories 3 and 4 are:-

3. Comprehension involving capacity to understand problems, to translate symbolic forms, to follow and extend reasoning.

4. Application of appropriate concepts in unfamiliar mathematical situation.

Hence the development of Bloom's Taxonomy of educational objectives has opened up a new area of study in education in general and Mathematics education and evaluation in particular. In the next section we look at evaluation procedure. That is the process and the instruments for evaluation.

2.4 Evaluation Procedures.

Evaluation process involves the use of a wide variety of procedures. The procedures may be classified and described in accordance with the frame of reference used. Gronlund (1981) classifies the evaluation procedures into maximum versus typical performances, and criterion-referenced versus norm-referenced measurement.

On the basis of measurement Cronbach (1970) classifies evaluation procedures into two main categories: measures of maximum performance and measures of typical

performance. The first classification concerns those procedures used to determine one's abilities. That is, how well one performs when motivated to put forth his or her best effort. Here aptitude and achievement tests are included in this category; Gronlund (1981). The second classification of procedures reflects a person's typical behaviour. That is how an individual performs under normal or routine situations. The evaluation result here indicate what an individual will do rather than what he can do. The use of personality appraisal methods is employed to evaluate a students'interest, attitudes and various aspect of personal-social adjustement; Gronlund (1981). The distinction between ability and typical behaviour can be observed when a student with considerable abilities to do more mathematics shows little or no interest in the pursuit of mathematics.

Gronlund (1981) observes that evaluation techniques used by classroom teachers provide a convenient frame-work for describing evaluation procedures. Airasian and Madaus (1972) report that one of such classification follows the sequence in which evaluation procedures are likely to be used in classroom instructions. These classifications are related to their role in teaching in the following way:-

1. Evaluation of pupil entry performance in a sequence of instruction (placement evaluation).

2. Evaluation of pupil learning progress during instruction (formative evaluation)

3. Evaluation of pupil learning difficulties during instruction (diagnostic evaluation)

4. Evaluation of pupil achievement at the end of instruction (Summative evaluation)

Gronlund (1981; 17).

Another classification of evaluation procedures is in terms of how the results are interpreted. Here again there are two basic ways of interpreting pupil performances in an evaluation technique. One of them is to specifically describe one's performance in an evaluation technique in relation to the course content. The second is to describe one's performance in relation to his classmate. The first type of interpretation is known as Criterion-referenced which according to Popham (1981) is used to ascertain an individual's status with respect to a defined behaviour domain; the second is called Norm-referenced which Popham describes as ascertaining an individual's status with respect to the performance of other individual in the group.

Gronlund (1981) notes that both interpretationsare useful. He explains that criterion-referenced interpretation enablesus to described what an individual can do, without reference to the performance of others. For example "Can achieve 56% of the instructional objectives in his geometry class". Norm-reference interpretation

on the other hand, enables us to determine how an individual's performance compares with others. The comparison can be on local basis, state basis or national basis depending on the intention of the evaluation results. For example, using state norm, we might describe a student's performance on a mathematic test as exceeding that of 66 percent of a state sample of secondary three students. In norm-referenced interpretation we simply describe what percent of the students in the norm group was surpassed rather than what percentage of geometry content was mastered.

Most teacher-made and standardized tests of achievement are norm-referenced. They are designed to rank students in order of achievement from high to low. This enables us to make sound decisions, based on relative achievement, for selection, grading and grouping students. The key feature for constructing norm-referenced tests is the selection of items that give a wide range of scores. This is carried out by eliminating those items that all students are likely to answer correctly and by including items at around the 50 per cent level of difficulty; Gronlund (1981).

On the other hand since the criterion-referenced tests are not concerned with the relative achievement of students, item difficulty level and descriminating power are not considered as criteria in items selection.

Rather they are selected on the basis of how well they reflect the specific learning activities being assessed. If the learning activities are difficult the items should be difficult. If the learning activities are easy the items should be easy. Glaser and Nitko (1971) look at criterion-referenced tests within the context of classroom instruction as one that is deliberately constructed to yield measurements that are directly interpretable in terms of a specified domain of instructionally revelant tasks.

2.5 Evaluation Instruments.

In the discussion above we imply the use of various techniques in evaluating students progress. We have also observed that when we evaluate students progress we make use of both quantitative and qualitative descriptions about the students' performance. Evaluation techniques can be said to be using three main instruments. These are measurement data which give the quantitative descriptions of the students' performances, the non-measurement descriptions which give the qualitative explanations of the students' work; and value judgement which give the worth of the results so obtained.

Brown (1976:8) after careful consideration of the definitions of measurement given by Guilford (1954) and Stevens (1951) restates the definition of measurement as:-

.....the assignment of numerals
to behaviour according to rules .

This implies that the characteristics of measurement involves the use of certain procedures or operations in accordance to specific rules. It involves the assignment of numerical values to a person's performance. It is clear from this definition that the numerical values should be expressed on a well defined scale. Measurement is limited to quantitative descriptions of the students performances. It answers the question, How much? that is, the results of measurement are always expressed in numbers. For example "Abdu scores 64% in his mathematics sessional examination". It neither include qualitative descriptions (Non-measurement) nor does it imply judgements concerning the worth or value of the results obtained. Thus measurement data provide a continuum in the assessment of some dimension in our progress. The goal of Psychological and Educational Measurement is to place an individual at proper place on this continuum.

It should, however, be noted that, when we interpret a student's performances, we usually assign some value or worth on it in terms of standard for his class (non-measurement). Be it the student's educational and/or vocational plans or some other basis for making value judgement. In this case we are no

longer restricting ourselves to "measurement" we are now "evaluating" his abilities or his progress; Adams (1966).

Evaluation always involves value judgement for desirability of the results. It may include either quantitative or qualitative descriptions of students progress or both. Thus answers the question, How good?

Gronlund (1981) illustrates the comprehensive nature of evaluation process:-

EVALUATION: Quantitative descriptions (Measurement)
and/or
Qualitative descriptions (Non-measurement)
plus
Value judgement.

It is therefore clear that evaluation may or may not be based on measurement. If it does, it goes beyond the simple quantitative descriptions.

Evaluators and classroom teachers use various instruments to obtain measurement data. They include among others;

Classroom assignment or home work
Project
Observations
Tests
Questionaires
etc.

Perhaps the most widely and frequently used measurement instrument by the teachers and evaluators are tests. This may be due to the fact that they give a good measure of the students achievement and progress

of instructional objectives set forth by the teachers and authorities within specified domain. In addition evaluators may prefer using tests to measure student's achievement because they are much more objective to prepare, administer and score than most of the other measurement instruments.

Test as a measurement instrument has been defined by a representative of the American Psychological Association as:-

> .....nothing more than careful observations of actual performance under standard condition.
>
> Adams (1966;5)

It can be infered by "careful observation" that the method of sampling the performance of individual and recording of it must be systematic. It must be objective enough so that comparable results can be obtained by different observers. The definition gives the concept of measurement. We record data on a sampling of performance under standard condition. Because testing is the main theme of this study, it is hereby discussed in more detailed below.

2.6 Types of Test.

Before discussing the categories of tests, it will be very useful to reconsider other definitions of tests so that a better insight of what they are may be gained. Gronbach (1970) gives a broad definition of a test to

cover testing in Psychology. education and industry.

To him:

> test is a systematic procedure for observing a persons behaviour and describing it with the aid of a numerical scale or a category-system .

Cronbach uses numerical scale here to describe a person scoring 15 sums out of 21 in Mathematics. While category-system is used when he is said to be an A student. This definition includes reports on personality through questionaire, methods of observing social behaviour, apparatus tests measuring co-ordination, and even systematic records of output on a production line. In this Cronbach looks at test as synonimous to evaluation.

Brown (1976;7) in his definition of a test states that:

> a test is a systematic procedure for measuring a sample of behaviour .

A careful analysis of the various components of this definition reveals a systematic construction, administration and scoring of items in accordance to specified rules. By "systematic procedures", it implies that the test items are selected to fit particular specifications. It implies that all persons take the same or equivalent test items and the directions and the time limits are the same for all person taking the test. The standard procedures of scoring the test minimizes the possible

influence of individual differences in scores for personal and situational veriables. It enables us to compare directly scores of different persons taking the same or equivalent items. This implies that if two persons take different items or different amount of time to complete the same test, their results can not be compared.

Another important term in this definition is "behaviour". The behaviour manifested on the test will only provide useful information if it adequately shows the construct being measured. Inferences from the test scores will therefore be in error if it does not adequately reflect the underlying characteristic for which the test is administered.

It should also be noted that a test contains only a "sample" of all possible items. Brown (1976) notes that no test can be so comprehensive enough to include all possible items that may be developed to measure the behaviour domain. Thus any test items should only be considered as sample of all possible items. And since a test contains only a sample of all possible items, we must assure that our test represents a sample of all possible items. If a person takes two different tests within the same domain he must obtain the same score.

Having clarified the concept of test, we can now ask ourselves the question, How many types of test do we have?

Tests can be classified into many categories depending on the intentions for which we are testing, the procedure followed in giving the test, the test content and the method of scoring the items.

The classification under intentions may follow the sequence in which the evaluation approaches are likely to be employed in the classroom instruction; Airasian et al (1974). The intentions are related to the role of teaching in determining the level of knowledge acquired from the study of a specified area. That is the level of achievement within a specific domain. The intention may also be to diagnose the students learning difficulties which will provide information for method of better performance, which Gronlund (1981) refers as <u>diagnostic evaluation</u>.

Another intention for which we administer tests is to assess students entry performance in a sequence of classroom instruction which provides necessary information for placement of student at the appropriate level of instruction. We may also give test to predict future performance of the students.

The classification for procedure determines the methods used in giving the test. It can be paper-and-pencil tests which require the students to write the answers to the questions on paper; OR performance tests which usually require the students to use

physical objects to demonstrate the amount of learning achieved by applying physical and motor skills. It can be group tests which are administered to individuals at the same time; or individual tests which are administered individually. Another classification under such procedural differences is speed and power tests. In a speed test the problems presented before the candidates are of approximately the same level of difficulty; with very limited time for administration such that no one or a few of the candidates can complete the tasks. The scores reflects the candidate's speed in performing the tasks presented. In power test almost all candidates are given sufficient time to finish the test. The problems are arranged in order of difficulty and the candidate's scores reflects his accuracy and the level of difficulty at which the candidate can successfully perform; Adams (1966).

The content classification of tests tells us about tests of different course areas. For example, test of ability in various courses; Geography, History, Mathematics or intelligence. It can be very broad like tests on course units or it can be more specified like tests on simulteneous linear equations or on immediate courses of Nigerian Civil War.

The classification under the method of scoring will select the appropriate item types. This can be objective test items or essay test items.

The objective test items are of different kinds. Perhaps the most commonly used ones are multiple-choice, true-false matching and short-answer type.

In multiple-choice items the candidate is required to select the correct answer from among a certain number of choices, say, four. It is widely applicable to various tasks performance, though somewhat more difficult to prepare than any other items types. Its main advantages are the objectivity in scoring and high reliability. Theor tically, multiple-choice test, though practically verified, is expected to show much more reliability than most other item types; Ebel (1979). However, critics of multiple-choice allage that the objectivity of the items is inevitably superficial, ambiguous and the responses are subject to guessing.

The true-false items require the candidate to either accept or reject a proposition. They are best suited for assessing knowledge that are unequivocally true or false. For example, "any four sided figure is a square". The candidate here has a 50 percent chance of getting the answer correct by guessing. Thorndike and Hagen (1977) observe that the format yields less accurate information about a student per item than do other forms of item, and many more items are needed to provide an equally precise appraisal of each candidate's competence. They therefore suggest that true-false tests

should have a relatively large number of items, and there should be time enough for all candidates to read and response to all of them.

The matching item traditionally consists of two paralled columns, with each contained word, symbol or number being matched to a word, sentence or phrase in the other. The column items for which a match is sought are called premises and the column items for which the selection is made are called responses. The directions provide some explanation of the basis for matching responses to premises. Usually the items in the column responses exceed the items in the column premises and candidate is expected to use any of the item in responses once, more than once or not at all. The typical matching item is limited to the assessment of factual information based on simple associations. Gronlund (1981) observes that the major advantage of matching item is its compact form, which makes it possible to measure a large amount of related factual material in a relatively short time. He also maintains that its major limitations are its restriction to the measurement of factual information based on rote learning, and its high susceptibility to the presence of irrelevant clues.

The short-answer items require the candidate to supply a word, phrase, number or symbol. They are suitable for measuring a wide variety of relatively

simple learning outcomes. Gronlund (1981) maintains that except for the problem-solving outcomes measured in Mathematics and Science, the short-answer items are used most exclusively to measure the recall of memorized information. They tend to be used very widely in classroom tests and they are easy to prepare. Ebel (1979) observes that they serve a very useful function especially in the Primary Schools and Junior Classes of Secondary School where the formation and development of concepts and vocabulary is most emphasized. They have the advantage of requiring the candidate to think and supply the answer.

The essay test items allow the candidate some freedom of responses. Gronlund (1981) classifies them into extended response type, where the candidate is given almost complete freedom in making his or her response, and the restricted response type, where some limitations are placed on the nature, length or organization of the response. Examples of these types are:-

1. Five litres of groundnut oil and three litres of palm oil cost ₦14.00 altogether. One litre of groundnut oil cost 48k more than a litre of palm oil. Find the cost of each per litre. (extended response)

2. Solve the following simultaneous equation by substitution:-

$3x - y = 1.$

$2x + 3y = 5.$ (Restricted response)

The extended response type permits the candidate to decide which facts he or she thinks is most pertinent, to select his or her own method of organization, and to write a comprehensive answer in his or her own way. Such items tend to reveal the candidate's ability to evaluate communications, relate them in a coherent manner, and express them clearly. Its major limitations lie with its inefficiency for measuring knowledge of factual material, and the scoring is difficult and subjective.

The restricted response type is more efficient for measuring factual knowledge and reduce the difficulty somehow, of scoring. But the highly structured task of the item makes it less effective as a measure of the ability to select, organize and integrate ideas which is one of the unique purpose serve by essay items.

On the other hand Brown (1976) classifies maximal performance tests into three broad categories; (1) the tests that shows whether a student has the necessary skills and abilities to succeed in further work in an area, that he calls aptitude tests. (2) the tests that measure a student's skills in an area, that he calls ability tests (3) the tests that measure educational

or training programme outcomes that he refers to as the achievement tests.

Aptitude tests according to Brown's analysis, measure the results of general and incidental learning experiences and its frame of reference is towards the future. In other words aptitude tests measure the individual's total life experience and the aim of testing is to predict what can be learned in the future. They indicate the ability to acquire certain behaviours or skills given appropriate opportunity.

The most well known aptitude tests are the general Intelligence - tests usually refered by the common man as "IQ tests". These are tests specially prepared to measure a broad and overall intellectual functioning of a person. Individually administered tests are the original and most important procedure of this type, though recently group intelligence tests have been developed; Brown (1976). The individual intelligence tests are both verbal and performance tests. While the group intelligence tests are, in many ways, similar to the verbal portions of the individually administered tests, though restricted to measuring skills that can be tested by paper-and-pencil format. The ability tests on the other hand show one's power to perform a certain task. These types of tests differ from the aptitude tests which indicate the power to learn to perform a

certain task. That is to say the ability tests refer to a current state and aptitude to a future state. The ability tests generally cover broad areas. They measure learning done in a variety of settings, measure skills other than academic ones and often involve performance testing.

Brown (1976) also considers achievement tests as measuring learning experiences in a relatively restricted learning situation such as a classroom or training programme. Their frame of reference is on the present or past, that is on what has been learnt. They are meant to measure the knowledge and skills on a certain domain. The content may be as narrow as one day's work (for example addition of whole numbers with carrying) or as broad as many years' of study (such as Primary School Mathematics or Senior Secondary Mathematics). In any case we are attempting to measure what a subject can do within a specified given time. Our main interest, however, is to find out what has been learnt as a result of a particular learning opportunities offered. Because of the fact that we can not measure every bit of knowledge and possible skill, the test items are, normally, only a sample of all possible items within the given domain.

Achievement tests can be on the basis of content which measure knowledge of Chemistry, Mathematics, Nigerian History, Physics, Biology and so forth. In

each of the subject area the achievement tests can serve as a survey which gives an estimate of a subject's overall level of performance in a content area; or as a diagnostic measure which, in contrast to survey, attempts to assess the relative weakness and strength in important component skills and thus help to determine the areas where an individual needs remedial work; or as readiness indication of an individual's possession of the skills needed to learn the material at the next higher level.

Achievement tests can also be made on the basis of the mode of response used. They can be paper-and-pencil tests or performance tests. Paper-and-pencil tests can also be divided into two main classes. Recall and recognition. On recall tests the subject is to make his answers. (For example, essay and completion items) while on recognition tests the subject is to choose between alternatives (for example, multiple-choice and true-false). The performance tests require the subject to demonstrate his skills (for example, dissecting a frog or operating an engine).

Another major distinction in achievement tests is between standardized tests of achievement and teacher-made tests. The teacher-made tests are constructed by the classroom teacher, or possibly by a team of many teachers in the same school. The content area of this

type of tests is more restricted to what has been taught in the class or school. This provides the school with the opportunity to cover the content area taught in detail in the test and the scores can be compare with reference to the subject's class-mates. The standardized tests, on the other hand are usually constructed by test publishers and are prepared such that they can be used in different schools. This provides a broad coverage of the content and includes materials covered in many schools. The coverage is ascertained by consulting widely used textbooks and seeking the opinion of curriculum experts, thus obtaining norms that are generally national in scope. Because this study is on standardized tests, the researcher will now concentrate on their development in the next sections.

2.7 Standardized Tests.

The Dictionary of Education by Good (1945) defines a standardized test as:-

> a test for which content have been selected and checked empirically, for which norms have been established, for which uniform methods of administering and scoring have been developed, and for which may be scored with a relatively high degree of objectivity.

The processes of standardization according to Good involve the establishment of a fixed scoring procedure of the test as well as adequate descriptions of the age, grade, race or other norms of the testees.

Cronbach (1970) defines standardized test as:

> ...One in which the procedure, apparatus, and scoring have been fixed so that precisely the same testing procedures can be followed at different times and places.

Cronbach further explains that some tests do provide tables of norms describing scores that are usually obtained by the representative subjects. Such tests with the norms tables are sometimes called "standardized test", and the process of gethering norm data is called "standardization". He however, cautions that a test may have a table of norms even though its procedures are not clearly specified, and a test with well stated procedures may not have norms.

Thorndike and Hagen (1977) look at the term "standardized test" as implying the availability of nomative data, and the publishers of such tests provide a range of normative information to help give meaning to the test scores.

They go further to enumerate the characteristics of standardized test based on their definition as follows:

> Normative data provide for a particular type of meaning, in that they emphasize comparison with others. They permit statements about how Hilary is doing in relation to others of her age, or others of her grade, or others from the same type of school system in which she is enrolled. They permit statements about

> how Hilary is doing in Maths relative to her performance in reading or in relation to her performance on a test of scholastic aptitude. They permit statements about how Hilary's school is doing in comparision with schools the country over, or school of a particular type, or schools in a particular type of community.

In essence the primary objective of standardized tests is to compare an individual or group of individuals with others within the same groups in different aspects of task performance.

The researcher considers "standardized tests" as tests which contain fixed set of items that are carefully and systematically prepared, with uniform directions for administration and scoring based on norms that are clearly defined as the reference group.

The "fixed set of items" refers to the standard content under consideration and the "uniform directions" provide for uniformity in administering and scoring conditions of the tests. The "norms" allows for interpretation of the test results that is meaningful within the group under reference. In essence the individual's test performance can be compared with that of others to provide for interpretation in a well-defined group, refers to as the "norm-reference interpretation". This kind of interpretation is considered as the most

distinctiwe characteristics of standardized tests. Standardized tests are usually called "published tests".

There are quite a number of published tests for use in schools especially in the developed countries. But the ones that are of concern to the classroom teachers can be classified into two as: achievement tests and aptitude tests; Graunlund (1981). It should, however, be noted that measures in the affective domain such as appreciation, attitude and interest are also of concern to the classroom teacher as the researcher indicated earlier. For the purpose of this study the researcher is only interested in the cognative achievement tests. He will therefore only consider the procedures, the functions and studies on standardized achievement tests in cognitive domain in the rest of this chapter.

2.8 Theoretical Aspect of Standardization Process of a Test.

Wallace and Larsen (1978) observe that standardized tests have been in educational settings for a long time. Adams and Torgerson (1966) report that when Rice in 1894 was faced with considerable pressure to bring into the curriculum, such new practical subjects as manual training and home economics; and also with considerable opposition on the part of educators who thought that it was not possible to teach all subjects in the curriculum, Rice decided to administer uniform tests in

spelling in a number of schools as a step towards studying scientifically the effectiveness of instruction under different time alloted to each subject. The next achievement tests were developed and published in 1908 which was refered to as Stone Reasoning Test in Arithmetic. Thorndike Handwriting Scale was published in 1909. Since then E.L. Thorndike has made a number and significant contributions in published test that he is generally considered to be the father of educational-measurement movement; Adam & Torgerson (1966).

Hammill (1971) observes that standardized tests are usually administered by specially trained persons in settings other than classroom. He goes further to suggest that:

> ...the information acquired is of a decidedly quantitative nature and tend to compare a given child's performance with national or reginal normative data. The results, therefore, are often reported in terms of quotients, scaled scores, grade equivalents or percentiles .
>
> Wallace & Larsen (1978; 19)

Although standardization procedure may vary, depending on the nature of the test, ETS (1965) suggests a typical sequence of construction and standardization of a test as:-

Planning the test;
Writing the items;
Pretesting items;
Preparing the final form;
Collecting the reliability and validity evidence; and
Developing normative and interpretive material.

Brown (1976; 274).

We now consider this sequence.

(a) Planning the Test.

In planning an achievement test, the test constructor is faced with the problem of thorough analysis of the content of the curriculum and the objectives set fourth by the curriculum. Ebel (1979) recommends that a good test is the one that set explicit specifications that indicate:-

1. forms of test items to be used
2. number of items of each form
3. kinds of task the items will present
4. number of tasks of each kind
5. areas of content to be sampled
6. number of items in each area
7. level and distribution of item difficulty.

When this plan is outlined the test constructor then prepares, as part of the test specifications, a two-way grid, sometimes called a test blue-print. In this kind of chart the tester relates the content topics with the objectives already set forth in the curriculum. He/she indicates the number of questions and percentage of the question in each topic and skill

in the blue-print, taking into consideration all the steps recommended by Ebel-above. It is, then that the tester begins to write the items as in the words of Hill (1972) that:

> The paper setter should make a blue-print a chart of specifications before he begins to write questions.

(b) Writing The Items.

The blue-print guides the test setter to write the appropriate item type. These could be either the objective or essay type of items or both. The items are then reviewed by other subject matter and testing specialists. The review usually covers content face validity and accuracy, grammatical construction and possible bias in the items; Brown (1983). The proposed items are edited revised and rewritten, reviewed again until acceptable items are attained. When items are written by testing specialists, they are independantly reviewed by the subject matter experts to give the items the kind of flavour that the educators think important; Brown (1979).

It should however be noted that standardized tests generally utilize alternative-choice items for reasons of adequate sampling of content, speed and accuracy in scoring and higher reliability. Multiple-choice items are often used among the alternative-choice items for its advantage over the

others. More items than will be required are constructed. The number of items required will then be selected from this pool, using the pretest data to identify the items with the acceptable characteristics; Brown (1979).

(c) Pretesting Items.

After writing the appropriate items and the editing and reviewing, the items are then arranged into test forms (four or five) for preliminary administration to a large sample of candidates who are similar to those who will take the final form of the test. Relevant items statistics are computed, on the basis of which some items are eliminated, others revised while others are retained in their original form. If possible further tryout and analysis are conducted. The best items are then combined into the number of test forms required in the final test.

At pretesting stage the testing conditions must be specified. The procedures for administering and scoring the test must be developed. A proposed time limits for the test must go along with the directions for both the test takers and test administrators. The data obtained from the pretesting will then guide the test setter to set the exact time limits for the final test which should be determined by practical considerations of the amount of time available for

testing the number of items that will be needed for an adequate level of reliability. Answer sheets, scoring keys as well as the procedures to ensure accurate and reliable scoring must then be developed.

(d) Preparing The Final Form(s).

The final form(s) of the test(s) is/are administered to a very large sample of the pupulation. The items so assembled in each of the forms must be balanced to meet several conditions in order to produce tests with required characteristics. They should represent the various skills and/or content categories as enumerated in the test plan; the number of items must be appropriate for the time limit; the item difficulty level and discriminating power (item analysis) must be determined; a wide distribution of scores must be obtained; the various forms of the test must be equated to each other; and finally, the items are reviewed by subject matter and test construction specialists; Brown (1979).

i. Computation of Item Difficulty:- An item difficulty index indicate the percentage of candidates who got the item right. The computation of the item ease is done by the formula:-

$$P = \frac{R}{T} \times 100$$

where P = percentage of candiates who got the item right.

R = the number of candidates who got the item right; and

T = the total number of candidates who attempted the item.

ii. Computation of item Discriminating Power:-
The discriminating power of an item provides an index of how well an item differentiates between candidates who have greater and less amount of knowledge of the specific content. That is it discriminates between candidates with high and low achievement. The candidates are divided into upper and lower groups. The upper group are the candidates who scored high in the test and the lower group are the candidates who score low in the test. Usually the top 27% and the low 27% are considered for the computation of discrimiting index discregarding the middle 46% of the candidates. The item discriminating index can then be obtained by subtracting the number of candidates in the lower group who got the item right (refered to as Rl) from

the number of candidates in the upper group who got the item right (refered to as Ru) and dividing by one half of the total number of candidates included in the item analysis (refered to as $\frac{1}{2}T$); Grounlund (1981). Thus the formula is written as:-

$$D = \frac{Ru-Rl}{\frac{1}{2}T}$$

(e) Psychometric Qualities.

The determination of the psychometric qualities of the test is a continual process. Even though some characteristics e.g. time limits, difficulty level, distribution of scores etc. can be determined from the pretest data, candidates' response to the completed test may differ from their response to the pretests. Thus analysis of the completed forms of the test are needed; Brown (1979). Here we will be interested in the average score for various groups as well as the distribution of scores to determine the consistency of the test scores. Hence Brown (1983 & 1979) comments that:-

> The consistency of the test scores must be determined, particularly equivalent form reliability if there is more than one form of the test. The validity of the test - always content validity, but often also criterion - related-must be established. Normative data must be collected.

Brown goes further to suggest that even though the process of test development is designed to ensure that the test will be adequate in these respects; empirical studies are needed to confirm the effectiveness of the test construction process.

i. Reliability:- Test reliability refers to the consistency and accuracy of the test. Mukherjee (1978) explains that a reliable test is the one which on repeated measure gives the same scores of the same candidate or candidates. Anastasi (1959) considers the reliability of a test as:-

> ...the consistancy of the scores obtained by the same individuals on different occassions or with different sets of equivalent forms..

There are various methods of estimating reliability coefficient of a test. But some are of only limited usefulness with standardized achievement tests. Perhaps the most effective methods that could be applied to standardized achievement tests are the measures of equivalence and split half. Of the two methods that could be used, needless to say, whenever there are more than one form of the test, evidence of equivalent forms reliability is essential. Equivalent forms reliability, is based on the equivalence of scores on two tests that are intended to be substitutable measures of the same thing; Tuckman (1975)

The calculation of the equivalent forms reliability can be conducted by Pearson product-moment correlation coefficient using the formula:-

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{\left[N\Sigma X^2 - (\Sigma X)^2\right]\left[N\Sigma Y^2 - (\Sigma Y)^2\right]}}$$

Ferguson (1981; 113).

Where r denotes the product-moment correlation coefficient, N the number of observations; ΣXY, the sum of the product of the paired observation, ΣX^2 and ΣY^2 the sum of the squared observation, and $(\Sigma X)^2$ and $(\Sigma Y)^2$ the sum of the observations squared.

The Coefficient of equivalence between scores on parallel forms of the test, administered in close succession can also be calculated according to Brown (1979) by using the formula:-

$$r_{(A)(B)} = \frac{\Sigma(X_A)(X_B)/N - (\bar{X}_A)(\bar{X}_B).}{(S_A)(S_B)}$$

Brown (1979; 70).

On the other hand the split-half reliability is a method of obtaining consistency estimate of single test that is split into two reasonable equivalent halves usually odd and even numbered items. As with the coefficient of equivalence the scores of the two

halves are regarded as the results of two sets of tests. The correlation between these two scores gives an estimate of reliability. It should be understood that, the scores whose reliability is determined by this method have been obtained from half-length tests. To get the consistency of the total test, it is necessary to correct the half-test correlation to the expected full-length value. This is done using the Spearman Brown formula:-

$$r_2 = \frac{2r_s}{r_s + 1}$$

Ebel (1979; 278).

Where r_2 is the reliability of the total test, and r_s the reliability of the half-length test.

ii. Validity:- The validity of a test refers to its value for judging the candidates on what it intends to measure. For a test to be valid it must be reliable and relevant. As Marshall & Hales (1971) maintain, that:

> if a test reliably measures what is relevant, it is said to be valid. If a test is not reliable, it cannot be valid since an unreliable test provides useless data.

Validity has been classified into so many categories. The ones of interest to the standardized achievement test developers are the content and

criterion-related validity. The essence of content validity is to determine the extent to which the test items are the representative sample of all the possible items that measure the given curricular area. (Mostly used in measuring achievement tests). For example, a mathematics achievement test in algebra should sample an individual's knowledge of the total domain of algebra course covered before the test.

Criterion-related validity is used to predict candidates'future performance or to estimate current performance in accordance to some significant variable; that is the criterion. To get the measure of the validity of the test, an index is obtained by correlating two sets of scores from the same population. This is done by comparing one test with another related field. For example, the predictive validity of mathematics attainment may be compared with regard to intelligence.

(f) Norms: perhaps the most distinguishing features of a standardized test is its ease of interpretation because of the provision of norms. Stanly & Hopkins (1972) view norms as:-

> ...merely transformations of the raw scores to a more meaningful scale derived from the performance of large numbers of pupils distributed over wide geographical areas and representing various types of schools .

The interpretation of raw scores will depend upon the various norms in comparing the performance of the different individuals in the norms or groups under reference. To this Remmers, Gange and Rummel (1963) comment that norms are:-

> The levels of performances on a test attained by defined groups of pupils. They differ from standards in that they describe what is rather than what should be..

Brown (1983) considers scores on standardized achievement tests as either norm-referenced or content-referenced. He further explains that the norm-referenced defines the groups on the basis of certain variables such as grade level, geographical area, type of school, socioeconomic level of the community or type of curricula. These groups make use of any of the several types of scores such as percentile ranks, standard scores or grade equivalent scores. The content-referenced tests on the other hand is set to test the proficiency standard. The scores tell us how well the group has understood the materials taught and the degree of retention of the learned materials by the examinees.

Ebel (1979) reports that for standardized achievement tests, the norms are usually in the form of table of equivalents between raw score and some type of derived scores, such as: grade equivalent scores, age equivalent scores, standard scores or percentile

ranks. For example the grade equivalent scores compare the students performance with their counterpart in the same class or grade level. The age equivalent scores indicate the level of performance typical to a given chronological age.

The standard scores on the other hand express candidate scores in terms of standard deviation either above or below the mean of the distribution. For example if X represent a candidates raw score on a given test, $\bar{X}$ the mean score on that test in the reference group and S the standard deviation of the scores in the group then the candidate's standard score denoted by Z is given by the formula:- $Z = \frac{X - \bar{X}}{S}$.

A linear transformation is sually done to avoid the decimal and nagative numbers in the standard score. This transformation demote a standard T scores given by the formula:-

$$T = 50 + 10z$$

This formula transforms the raw score to a scale with a mean of zero, and standard deviation of one. Thus if a candidates Z - score is 1.3, it means that the candidate is one point three standard deviation above the mean distribution of the group under reference. When a candidate Z - score is - 1.7, his score is one point seven standard deviation below the mean distribution of the reference group.

made of them. However the Indian and American cases give some ideas about the qualities expected of standardized tests.

Kulkarni (1968) reports on constructing equivalent achievement survey tests in India at an international conference of experts on educational testing. The test in mathematics was aimed to be used for different student populations in India. Two alternative forms tests were finally constructed and printed after try-out with four alternative forms in thirteen languages of India and administered to students from different states at three different grade levels. These are grades 4, 5 and 6 at the Primary level; 7, 8 and 9 at the middle level; 10 and 11 at high school level. On the basis of the data collected from Delhi Schools, the correlation between two groups means in a set of schools was computed. This was aimed at ascertaining the extent to which the two tests were parallel. At Primary level the reliabiligy coefficient of .64 was obtained while .80 and .89 were computed for middle and high levels respectively, when same students took both forms of the tests. On the other hand when the students were randomly divided into two equivalent groups and Form I test was given to one group and Form 2 to the other, the reliability Coefficient of .83, .87 and .86 were calculated for primary, middle and high school levels accordingly.

Nair and Soman (1974) prepared a standardized Mathematics achievement test for grade 9 in the State of Kerala. They use the syllabuses for grades 8 and 9

to construct the items to measure the 6 levels of Bloom's cognative domain. Two forms of the tests were finally constructed and administered to 650 grade 9 students.

The reliability coefficient was calculated by both split-half and test-retest methods. .81 and .79 respectively were obtained as the coefficients for 100 students. The validity coefficients calculated were .75 and .82 using the average of two terminal marks of 200 students and the Kerala University Test in Mathematics Achievement of the same students in the study respectively.

In another study in mathematics, Buch (1971) reported that Venugopal, N.K. designed standardized achievement test for grade 8 based on sex, locality and management of schools norms. 950 pupils were randomly selected based on stratified sampling. The maximum score of the test was 50 and the minimum was 5. The data of this study revealed that boys performance is higher than girls and the performance of private institutions was better than the Government Schools. Split-half method was used to determine the reliability coefficient of .90. The validity using teachers' rating was calculated as .89.

Buch (1971) reported that A.K. Basu prepared group test on problem solving ability in mathematics for grade 9. The test was standardized for Bengali speaking

children after completion of grade 8. The standardization sample was 982 students of grade 8. The highest score was 40% and the lowest 2%. Reliability was calculated by split-half method and yielded very good result. Percentile norms and age norms were calculated.

Standardized achievement tests were constructed for High School Classes in Utter Pradesh in Physics and Chemistry by R.K. Bountara as reported by Buch (1971). The tests were constructed from the study of grade 9 and 10 physics and chemistry syllabuses. Test-retest reliability coefficient was calculated for each subject. Physics coefficient was .932 and that of chemistry was .950. The rational equivalence method of calculating reliability indicate .84 for physics and .89 for chemistry. By split-half reliability the coefficient for physics was .79 and .79 for chemistry. The validity coefficient for physics was found to be .596 and .605 for chemistry.

Braswell (1978) reports on the mathematical portion of Scholastic Aptitude Test (SAT) which is designed to assist colleges in evaluating the academic prospects of students. Frequently the student's high school record and scores on the SAT are used by colleges in America to predict how well a student is likely to succeed in their programs. The items used on SAT are pretested at national administration dates on a

random sample of students similar to those who will take later editions of the test. The pretested results are analysed statistically for a representative sample of about 2000 students, and the resulting data are reviewed. The analysis involves the percent of testees who answered the item correctly, the number who selected each alternative, the mean total test score for the group selecting each of the answer choices, and the correlation between performance on the item and performance on the SAT. The items whose psychometric qualities are acceptable go into a large pool of classified items from which new editions of the SAT are assembled.

The SAT scores are reported on a scale which extends from 200 to 800. It would, however be noted that a score of 200 does not necessarily mean that a candidate has answered every item incorrectly nor does a score of 800 imply perfection. The score scale uses an arbitrary means of 500 and a standard deviation of 100. The students' performances in the test have, however, shown some decline in Mathematics achievements. For example in 1958/59 the mean performance of all the candidates taking the Mathematics portion of SAT was 498 with a standard deviation of 110, while in 1969/76 a mean of 488 and a standard deviation of 113 were reported and a lower mean of 471 and a standard deviation of 117 were recorded in 1976/77 results.

Ferguson and Schmeiser (1978) write on the Mathematics Usage Test of the American College Testing Program (ACT). The ACT is primarily designed for college planning. High Schools in America use the results to offer precollege guidance services and educational planning. The on campus use of ACT test data include selection, placement, course sectioning, advising and counseling, and educational planning. Each year, three new forms of the ACT Assessment tests are completed and administered. The items in each forms are completely new. They have never appeared in any previous form of the test.

The development of new ACT Test is done, first by preparing a table of specification or blue-print based on content, cognative skills emphasis and item format of the test. Second the item writing is done by able high school and college educators selected throughout the country. Third the ACT test specialists carefully review and evaluate the items according to explicit standards required. The tryout followed after the items have been assembled. Fourth statistical test analysis is conducted on the data collected. This analysis examines the statistical performance of each item in the test and the performance of the test as a whole to confirm that the test and its items functioned as expected. The scores on the new forms are then equated

with the standard scores scale so as to ensure that the new ACT test forms compare favourably with scores of the previous forms of the test.

The scores obtained by students on the Mathematics Usage Test are used to evaluate their educational achievement by comparing their scores with the scores of various other important reference groups. The comparision is done with the use of norms that are provided by ACT.

The trends of scores for all students who completed the ACT in each of the year show test-score decline in Mathematics achievement test results. For example the means scores of all the Mathematics Usage Test candidates in 1964/65 was 19.6 with standard deviation of 6.6, while in 1972/73 the mean was 18.8 with standard deviation of 7.2 and 17.1 and 7.7 were recorded as the mean and standard deviation respectively in the 1975/76.

ii. Studies in Nigeria.

Formal testing programmes in education started in Nigeria with the establishment of Western types of schools. Ikeotuonye (1982) reports Nduka (1964), Lewis (1965), and Fafunwa (1974) as commenting that the Missionary stations, and along with them the Mission Schools were first established at

Badagry and Abeokuta in 1842. Thereafter testing in education featured prominently in the formal education system in Nigeria. The development of Western Education in West Africa along with which was the increasing number of schools and rising number of school population necessitated the establishment of a body to be responsible for examining school candidates for placement and certification. Ohuche and Akeju (1977) reported that an ordinance to the establishment of this body was passed in Ghana in 1951 and later made effective in Nigeria, Sierra-Leone and The Gambia. The body known as West African Examination Council had its first meeting at Accra, Ghana, in 1953. Since then it has grown every year with several branches. For example in 1966 the Test Development and Research Office (TEDRO) of the Council was established. This branch of the Council is now responsible for the standardization and the qualities of all the tests administered by the council.

According to Ohuche and Akeju (1977), two categories of tests can safely be classified as "standardized" in West African Scene. These are the International Developed tests (I – D) consisting of twenty-one sub-tests and the several common entrance examinations to secondary schools; the school certificate and the general certificate of education examinations which are prepared and administered by TEDRO.

The I-D tests are Aptitute tests specifically designed for use in Africa. They were designed and studied by Schwarz. Ohuche & Akeju reported that Schwarz (1964) writes:

> During February - April, 1962 a sample including one of every twelve such students was tested on ten I-D tests. The 67 schools participating in this study represented all provinces of Nigeria and various types of schools ..... In July - August, a similar study was carried in the Eastern Nigeria Primary Schools. Here ..... were tested on twelve ID tests..... Next, data were compiled on students in the Nigerian Secondary (grammar) Schools. The students of fifteen Grammar schools from the North, East, West and the Federal Territory were tested on all the I-D tests used at this level. Altogether, the sample included 4,023 students representing the six classes of the Nigerian Grammer School System. Finally, an additional set of fourteen Northern Nigeria Primary Schools was tested on tests developed after the first study. ..... The entire sample includes 12, 388 students from 215 schools.
>
> Ohuche & Akeju (1977; 139-140).

Schwarz used a representative norming groups which he later sub-divided into several variables. These include sex: boys and girls separately, and education; early secondary, higher secondary, secondary leaving, and early post secondary. Ohuche and Akeju reported that he even attempted to use ethnic and geographical groupings.

His norms were reported in Stanine.

In his norms for similarities at the end of primary education Schwarz reported boys' and girls' performance in the tests. The pupils raw scores were compared with the Stanine norms. A score of 21 and above attrack the highest point of the category system and 7 and below attrack the lowest point for both boys and girls. While his norms for Verbal Analogies Test (VAT) with 3-4 years Secondary Education, he reported students' raw scores of 36 and above as the highest point in the system for both boys and girls. The lowest point was 12 and below for boys and 10 and below for girls.

The second category of tests of the West African Examination Council that can be considered as standardized according to Ohuche & Akeju are the School Certificate Examination and the GCE 'O' and 'A' Levels. These tests are attainment tests given to students at the end of Secondary Education and the private candidates who have the equivalent of Secondary Education at both

'O' and 'A' Levels. There are tests on some forty school subjects available. These examinations are only standardized to the extent that; the mark are assigned after a standardized fashion, the subject grades are given on the basis of the raw marks obtained, and the final certificate division are given according to fixed standards; Ohuche & Akeju (1977).

In essence the West African Examination Council tests on placement and certification fall short of standardized tests in that the standardization process is incomplete. There is often no norms on previous sample, the assignment of the grades are usually subjective and not fixed.

A validation study of the Differential Aptitute Tests (DAT) was carried out in Kaduna State by Ikeotuonye (1982) to determine the applicability of DAT in Kaduna State for the purpose of educational and vocational guidance and counselling. The tests were given to a stratified random sample of 400 Secondary School Students whose ages ranged from 13 to 23 years. The reliability coefficient observed from the study ranged from .34 to .94 and the standard errors of measurement from .33 to 2.86. The results indicate that DAT are reliable and can be applied appropriately for the purpose of guidance and counselling in Kaduna State.

An attempt was made by Garba (1984) to standardized some fifty test items in Integrated Science for use as leaving Test for Junior Secondary Three in Gongola State of Nigeria. The researcher used stratified random sample to come up with the final number of students to take the final test. 965 students were selected out of 26,335 students in form 3 in all the schools in the State. The items try-out was done with a stratified random sample of 364 students. After the item tryout the researcher undertook the item analysis of the 80 objectives items initially prepared. Item difficulty and item discriminating power were computed. The researcher then came up with 50 items which according to him satisfy the criteria for selecting the final items.

The final test was administered to 965 students in 20 schools all over the state. The consistency of the test was calculated by the use of both split-half and test-retest methods. The reliability coefficient of 0.634 was obtained by split-half using the results of fifty students and 0.811 by test-retest approach again using the results of fifty students who took both the preliminary and final tests. The validity index was found to be 0.796 by using the results of fifty students in the final test and their average marks for three terms school examinations.

The norms were obtained by the use of frequency tables for the various strata, such as total population, total boys, total girls, urban subjects, rural subjects etc; cummulative frequency table for the various strata and percentile norms for the various groups. Measures of central tendencies, measures of dispersion and skewness and kurtosis were also calculated.

Observation: It is hereby observed from Garba's studies that:-

(a) the process followed in selecting both the schools and the subjects for both try-out and final testing has not been thoroughly explained in the research methodology: This made it doubtful to justify the representativeness of the sample that took the test to represent all the forms 3 students in Gongola State . Moreover in standardization what is required is a large sample of the population rather than randomised sample. The students that took the test if compared with all the form 3 students (which the research claimed the test to be meant for) represented only about 3.66% of the pupulation. This percentage is certainly not good for standardization.

(b) The researcher has not indicated the research time table to enable assessment of the test as to whether it could be used as leaving test for Junior Secondary Three. However, judging from the content in the blue-print and the test item, the test mostly covered form II integrated science work. Therefore the test as it is could not be meant as leaving test for form III.

The researcher believes that there could be quite a number of attempts in standardization of tests in Nigeria, the reports of which are not easy to come by. This makes it difficult for him to include more of such studies in this section of the literature review. The ones reported above are so far the only ones he is able to lay his hand on. With this one may want to know what implications this kinds of studies have on Nigerian Education System.

Brown (1983) enumerated the uses of standardized achievement tests in instructional process, measuring learning out-comes, evaluation of instruction and research. Standardized achievement tests are used to provide information that can be used to improve instruction which will consequently improve students' achievement. The main functions here are to help the teacher plan his instruction well as his students performance would reflect the amount of efforts he puts

in his teaching. They help the teacher identify students having special needs who need extra attention as a results of their mental ability which may either be so high or so low. Standardized achievement can be very useful in selecting students for particular course or place them at the appropriate class or school typo. They can as well be very useful instruments for educational and vocational guidance. The data from standardized achievement tests can be used to evaluate the teachers, curricula, instructional materials, educational programmes and instructional method.

The New National Policy on Education in Nigeria emphasises the quality of instruction at all level of learning; the implementation of continuous assessment in schools which is guidance oriented; and the selection and placement of students at the appropriate level of education based on their needs and capabilities. One of the ways that Nigeria can evaluate and control the quality of instruction of every teacher is by administering standardized achievement tests base on National, State and Local norms. This will give the teacher, the students as well as the government, the opportunity to compare the students' as well as the school performance with the national, state or local norms. Hence both the teachers and the students are put at alert for effective teaching and learning activities.

Chapter 3

DESIGN OF THE STUDY.

3.1 The Population

It should be recalled that the main objective of this study was to construct and standardize some test items from the secondary 3 algebra course of the contemporary Junior Secondary School Mathematics Curriculum. The purpose of which is to be used for assessing algebraic achievement of all the Form 3 students in Kano State. These standardized achievement tests in algebra, aimed at helping the appropriate schools authorities check for the subjective judgement, that classroom teachers are likely to make in the implementation of Continuous Assessment system in the State. We have already noted that one of the problems identified by the National Steering Committee for the Implementation of Continuous Assessment in the country was the comparability of standard across schools. This study was aimed at ensuring getting the required standard, placement of students to the appropriate course in algebra in the senior secondary institutions and certificating the graduatesof the junior secondary schools in Kano State.

To achieve these purposes, the entire population of Form 3 students were considered. There were about 9113 students in 41 secondary institutions which had up

to form 3 in the state at the time this study was conducted in 1983. The number of schools include 32 Government Institutions, 7 Private Institutions and 2 Federal Government Colleges. The population to which this study would apply directly, should therefore be all the Form 3 students in Kano State. Whenthe tests are to be used for the purposes they are designed the results of this study would only be comparable directly with Kano State Form 3 Students in the Junior Secondary Schools. Their use in any other states may be for research purposes and validation exercises.

3.2 Sample.

The subjects of this study were Secondary School Students in Kano and Kaduna States. The subjects from Kaduna State were involved at the first try-out of the test items, the performances of whom are discussed in the next chapter. The subjects from Kano State included a sample of students who just started Form 4 for the second try-out and the students finishing Form 3 for the final testing. Both these groups were expected to have completed the algebra course in this study as presented in Appendix A.

The purpose of sampling is to achieve representativeness of the characteristics of the peremeters of the parent population. When random sampling is employed,

Table 3.1 Institutions Selected for the First Try-out in Zaria Local Government Area of Kaduna State.

S/NO	INSTITUTIONS	SEX	SAMPLE FORM 3
1.	Kufena College, Zaria	Boys	64
2.	Govt. Day Sec.Sch. Zaria	Mixed	62
3.	Barewa College, Zaria	Boys	79
4.	Comm. College, Zaria	Mixed	81
5.	Girls Sec. Sch. Soba	Girls	63
	TOTAL		349

The Second Try-out involved 6 Institutions which were randomly Selected based on stratified sampling. Table 3.2 below shows the Institutions involved.

Table 3.2 Institutions Selected for the Second Try-out in Kano State.

S/NO	INSTITUTIONS	SEX	SCH TYPE	POPULATION (FORM 4)
1.	Fed. Govt Coll. Kano	Mixed	FG	153
2.	Sc.Sec. Sch. D/Kudu	Boys	P	203
3.	Tech. College, Wudil	Boys	P	183
4.	Govt. Sec. Sch. Hadejia	Boys	SG	65
5.	Govt. Girls Sec.Sch. Kwa	Girls	SG	28
6.	Govt. Girls Sec.Sch.Kano	Girls	SG	27
	TOTAL			659

Key:- FG = Federal Government (Institution)

SG = State Government (Institutions)

P = Private (Institutions)

(ii) The Sample For The Final Testing: The final administration of the test involved fourteen institutions all over the state. Like in the Second Try-out the sampling employed was stratified methods to take care of the different types of institutions and sex. The number of institutions were to be fifteen but the researcher was unable to administer the tests in any of the two Federal Government Colleges. The reason being that at the time the tests were to be administered the two Colleges were busy involved in the interview for admission into Form 1 classes and their sessional examinations were on. When he requested for the tests to be taken after their sessional examinations, he was made to understand that the students would vacate the very day they finished the examinations. However, the Table below shows the Institutions selected for the final testing.

Table 3.3. Institutions Selected For The Final Testing

S/NO	INSTITUTIONS	SEX	SCH. TYPE	POPULATION (FORM 3)
1.	Govt. Girls' Sec. Sch. Kura	Girls	SG	196
2.	St. Louis Sec. Sch. Kano	Girls	P	72
3.	Govt. Girls' Sc. Sch. Kwa	Girls	SG	102
4.	Govt. Girls' Coll. Dala	Girls	SG	250
5.	Aminu Kano Commercial College, Kano	Mixed	P	366
6.	Govt. Sec. Sch. Gaya	Boys	SG	264
7.	Ahmadiyya Sec. Sch. Kano	Boys	P	67
8.	Govt. Sec. Sch. Gwarzo	Boys	SG	340
9.	Govt. Sec. Comm. School, Wudil	Boys	SG	175
10.	Govt. Sec. Sch. Dambatta	Boys	SG	215
11.	Govt. Girls' Sec. School, Kabo	Girls	SG	55
12.	Govt. Sec. Sch. Gwale	Boys	SG	334
13.	Science Sec. Sch. D/Tofa	Boys	P	232
14.	Rumfa College, Kano	Boys	SG	413
	TOTAL			3,081

3.3 Design Implementation.

The general research design of this study was the equivalent Form tests or parallel tests strictly following the standardization processes under taken by test publishers. The two, preliminary testings or try-out consisted of five sets of tests (included in Appendix D) which were meant to be parallel. The final testing comprised of two equivalent tests, in Appendix F,

formed from the five tests administered at the preliminary stages in Kaduna and Kano States for pilot study.

The selection of the final test items in the two tests was made on the bases of difficulty level, discriminating power as in Appendix B, content validity and objectives of the area under study as indicated by the curriculum.

The norms were established on the basis of grade at state level, school types and sex; that is boys and girls. The reliability estimate of the two tests, the percentile norms as well as standard score norms are shown statistically in chapter V. The method employed in the development of the tests and administration, scoring and interpretation of the test results will be discussed later in chapter IV.

Chapter 4.

DEVELOPMENT OF THE DATA COLLECTION INSTRUMENTS.

4.1 Introduction.

The main purpose of this study is to develop standardized achievement tests in secondary 3 algebra which would help the classroom teachers evaluate students performances in this section of mathematics as compared with their counterparts in other schools in Kano State. This chapter is concerned with the construction of the data gethering instruments, and their relisbility and validity.

Two strategies were used to develop the instruments for this study. First, test items were constructed by the researcher. Throughout his reading and discussion with educators and evaluators in Nigeria the researcher is yet to come across some information about standardized achievement tests in mathematics developed in Nigeria. by Nigerians. Most examining bodies in the country, like WAEC and NTI, stop at trial testing their test items before administration for the purpose of certification. Norms have never been developed in the strict sense of standardization of achievements by these examining bodies. In the interest of Continuous Assessement in schools,

two parallel tests were finally designed and standardized, following the procedure for test construction in standardization process.

Second, a one day workshop was prepared to help the researcher develop valid and reliable items as well as administration procedure. These are discussed below.

4.2 Items Constructed.

In developing the test items the researcher throughly analysed the curriculum in Appendix A. The content and the objectives were first specified. A two-way grid was then developed to relate the content with the behaviour outcomes expected at class 3 algebra course to form a model of the test plan. The main idea presented in the model of the study therefore related the algebra content of secondary 3 with the behaviour outcome developed by the researcher.

A look at the content reveals that three main topics are supposed to be treated. These are:-

1.0 Factorization

2.0 Equations

3.0 Variation

It is the researcher's views that from the curriculum, all cognitive behaviour manifestations presented by Bloom (1956) in his famous Taxonomy of

Educational Objectives, can be tested at this stage of the students' education. But he does not feel that they need to be presented in the same form discussed by Bloom. It should be noted that, Avital and Shettleworth (1968); Wilson (1971); Husen (1967) and Wood (1968) to mention a few came up with their own model of evaluating mathematics achievements. For developing the model of this study the following stages of cognitive behaviour have been identified:-

A.0 Knowledge

B.0 Mathematical Skills

C.0 Problem-solving.

A table was developed to guide the construction of the test items which would test the expected behaviour outcomes. The proportion of each as shown in Table 4.1 below was decided by the researcher based on the relative emphasis on the curriculum, the time that would likely be spent on teaching each topic and the length of the topic treatment in the textbooks written on the curriculum. It has been observed that more emphasis in term of length is given to solutions of equations which calls for acquisition of skills. Hence the largest proprotion given to equations and skills in this model. The proportions of other areas were also based on the same principles.

Table 4.1. Guideline For Constructing The Test Items For Standardized Achievement Test For Junior Secondary 3 Algebra.

	CONTENT	A.O.KNOWLEDGE	B.O.MATHEMATICAL SKILLS	C.O.PROBLEM SOLVING	TOTAL	PERCENTAGE
1.0	Factorization	8	2		10	37.04%
2.0	Equations		10	2	12	44.44%
3.0	Variation	1	2	2	5	18.52%
	Total	9	14	4	27	
	Percentage	33.33%	51.85%	14.82%		100%

Each of the five sets of tests administered in the two try-out as well as the two sets of tests at the final testing carried the above model. There were 27 items in each test as indicated by Table 4.1.

After a more critical look at the curriculum the above table has been expanded to provide for more information about the content and the expected behaviour outcomes.

The content area is expanded as follows:

1.0 Factorization

1.1 Factorization by grouping

1.2 Difference of two squares

1.3 Quadratic Polynomial

1.4 Application of Factorization to Arithmetic.

2.0 Equations

2.1 Solutions of Equation involving Fractions

2.2 Word Problems involving Fractions

2.3 Solution of Simultaneous Equations including Word problems.

2.4 Changing the subject of the formula.

3.0 Variation

3.1 Direct variation

3.2 Inverse variation

3.3 Partial variation

3.4 Joint variation.

The behaviour outcomes are expanded as follows:-

A.0 Knowledge

A.1 Knowledge of Terms

A.2 Knowledge of Concepts

A.3 Knowledge of Generalizations

A.4 Knowledge of Rules

B.0 Mathematical Skills

B.1 Ability to compute and manipulate with numbers

B.2 Ability to translate and Interpret Mathematical Problem

B.3 Ability to solve routine problem.

C.0 Problem-Solving

C.1 Ability to analyze data and Recognize patterns

C.2 Ability to solve Non-Routine Problems

C.3 Ability to Discover Relationships

C.4 Ability to critize procedure for solving Mathematical Problems.

A more detailed table of specification was then developed below to provide for more information about the test content and the expected behaviour outcome. Same principles used to develop Table 4.1 were applied in developing this table.

Table 4.2 Table of Specification for the Standardized Test for Junior Secondary 3 Algebra Test.

CONTENT	BEHAVIOUR											
	A.0. KNOWLEGE				B.0.MATHE. SKILLS			C.0. PROBLEM-SOLVING				
	A.1 Terms	A.2 Concepts	A.3 Generalizations	A.4 Rules	B.1 Computation & Manipulation	B.2 Translation & Interpretation	B.3 Routine Problems	C.1 Analysis of Data & Recognition of Pattern	C.2 Solution of Non-Routine Problems	C.3 Discovery of Relationship	C.4 Criticize Procedure	TOTAL
1.0 FACTORIZATION												
1.1 Factorization by Grouping		1	1	2								4
1.2 Difference of Two Squares			1	1								2
1.3 Quadratic Poly-nomials		1	1									2
1.4 Application of Fact. to Arith.							2					2
2.0 EQUATIONS												
2.1 Equation involving Fract.					2							2
2.2 World Problem									1			1
2.3 Simultaneous Equation						2	4	1				7
2.4 Changing the Subj. of the Formula						1	1					2
3.0 VARIATION												
3.1 Direct Variation	1					1				1		3
3.2 Inverse Variation											1	1
3.3 Partial Variation												
3.4 Joint Variation							1					1
TOTAL	1	2	3	3	2	4	8	1	1	1	1	27

The test-items were developed from books written for the New Curriculum under study. They include New General Maths for West Africa by Channon, Smith, Head, Ialejaiye and Macrae; Oxford Secondary Mathematics by Adelaja, Lassa, Oyelese, Paling and Wardle; Junior Secondary Mathematics by Okosi, Ferris and Busbridge; Macmillan Junior Secondary Mathematics by Ojo Ogam and Sambo; and New School Mathematics by Emma, Jahun, Chee and Sin. In particular the researcher used his experience in writing of books for the curriculum which was based on his classroom teaching to write the test-items. Originally, 212 short answer and essay items were constructed. The researcher is fully aware of the facts that most standardized tests are objectives types. The use of short answer and essay type items in this study is motivated by the fact that the model so developed could best be tested by these type of items. It is the researcher's convinction that the acquision of skills and problem-solving behaviour could best be assessed through the process employed by the students in solving problem. This will indicate to the teacher the areas where the students need some remedial work and the steps where the students missed the process. They will also take care of correction for guessing and show the actual effort put by the students in their attempt to

answer the question. Moreso in Mathematics the scoring of these types of items could be very objective by using a carefully prepared marking scheme.

A copy each of the items constructed was given to five teachers of Form 3 in some schools in Zaria and two lecturers in Ahmadu Bello University for content validation. It should be recalled that we have already noted in chapter 2 that in standardization what is always needed was content validity, though often also criterion-related validity may be needed, Brown (1976). For this study content validity was employed.

The teachers and the lecturers went through all the 212 items and passed their comments and suggestions. The comments and suggestions were incorporated and five sets of tests were developed each containing 27 questions. These were supposed to be parallel. In order to be sure of their face equivalent the tests were again given to the Secondary School teachers for their criticism. All the teachers accepted that the items were equally weighted for the content as well as difficulty level and standard for the topics and class under study.

4:3 First Try-out

The five sets of tests were administered to Form 3 students as indicated in Table 4.1. To help in checking the difficulty of the tests, the same sets of tests were

administered to Form 2 and Form 4 students. The tests were administered in close succession by the researcher with the help of some teachers in the various institutions between June and July, 1982. The scoring of the tests were done by the researcher and the results were very poor. There was no single item that was scored correctly by more than 42 students out of 349 students who took the test. The highest score in the test was 11%. Very few students scored between 2% to 11%. Most of the students scored 0%. To the researcher the results did not give anything to study. The results were therefore discarded. See Appendix B for some results of this try-out.

It was however observed during the test administration that there were some abnormalities in the conduct of the test and the students attitude towards the tests. One of the abnormalities was caused by the researcher's ignorance of the amount of teaching that was done to the students in all the schools. The researcher assumed that the students had had some three years of good mathematics teaching. But it was discovered through discussions with the students as well as the teachers that the students were never taught for one or two years out of their 3 years study in the schools. Another abnormality is that the students did not seem to be

serious about the test because they knew that it was not going to count on their reports for the session. The students were therefore not in examination mood.

In order to find out the real condition in the schools with regard the teaching, the content of algebra in secondary 3 and to control the abnormalities identified above, a one day workshop was conducted involving teachers at secondary schools level in Kano State.

4.4 The Workshop.

With the help of the Chief Inspector of Education in charge of Kano Educational Resource Centre a one day workshop was conducted on 9th October, 1982 at the Centre. Fifty nine (59) teachers of Form 3 and 4 all over the State and the State Inspector of Mathematics represented twenty nine Institutions. The workshop discussed among other things the algebra syllabus taught in Forms 3; the text-books used; the strength of mathematics teaching staff in schools in the state; the average number of period per week; and the test-items developed for the Kaduna State try-out.

During the workshop the teachers filled in the research information sheet which gave the researcher some ideas about the strength of mathematics teachers, their

qualification and experience as well as the number of periods taught in a week. The information collected helped in identifying the Institutions that could be involved in the study. (The Proforma in Appendix E.)

The topics under study in the new curriculum were presented to the teachers for their scrutiny. It was discovered that some topics were treated in Form 2 using the old syllabus while others were to be treated in Form 3. Some adjustments for the whole state in the old syllabus to fit in the new syllabus were suggested by the State Inspector for Mathematics which were accepted by all the schools. This was as a result of the discussion of the whole syllabus for secondary schools which prepared students for GCE/SC Examinations. It was discussed that all the algebra topics presented be spread to cover the teaching for the whole year of Form 3 in the State.

With regard to text-books it was discovered that most Institutions in the State used New General Mathematics for West Africa. It was then agreed that the book be used plus any other text-books that covered the topics in the new syllabus.

There was a thorough discussion on the 185 refined items from the original 212 items used in Kaduna State. Many suggestions were made by the participating teachers

for improving the items. These were incorporated and further review of the items was made.

Later the teachers' assistance was solicited in adding more items to the bank. Many teachers submitted their terminal and sessional tests given to their students to the researcher and promise to send some more items. A week later the researcher started to receive more items developed by the individual teachers from institution through the State Inspector. These were incorporated into the items already discussed and used in the second try-out.

4.5 Second Try-out.

On 26 November, 1982 five sets of tests were administered to 659 Form 4 students in six institutions as presented in Table 3.2. The tests were administered in close succession on the same day in all the institutions. The same five sets of tests were also administered to 66 Form 3 students in the same manner in 3 of the institutions at the same time. This was intended to check the difficulty level of the tests using the set of students who would be taking the test by the time they completed the course in June, 1983.

Each of the five tests included 27 items which were content validated by six experienced Form 3 and 4

teachers whose schools were not taking the tests and the State Inspector of Mathematics. The criteria of validation were course objective, content coverage, scope and standard of the tests. All teachers agreed that the items were within the experience of the students with the exception of the items on variation and changing the subject of the formula which were not treated in either forms 1, 2 or 3 in the old syllabus. However, it was ascertained that the items so developed in these sections were the types within the standard. The tests questions were the modified version of the items discussed at the workshop and the submission of the various teachers who attended the workshop.

(i) Administration:

Each of the five sets of tests was arranged according to varying level of complexity. They were then typed and duplicated and arranged in envelopes according to the number of students in each class for each school. The researcher went to each of the schools and submitted the examination papers. These included the question papers on which the students answered the questions, graph papers, a proforma requesting the teachers' comments about the examinations, direction for administering the tests and the marking

schemes. The marking schemes were throughly discussed with the teachers. The instruction to the students was also written at the cover for the question and answer. Papers. As mentioned earlier all the schools took the tests in the morning of 26 November, 1982.

(ii) The Results.

The teachers in the schools scored the tests strictly using the marking scheme. They later submitted all the test papers together with the students' results as requested in the instruction to the examiner. Four of the schools which took the tests reported that they used the same test results for their terminal reports. The students performances were later subjected to statistical analysis.

(iii) Statistical Analysis.

Since the five sets of tests were meant to be equivalent, the first analysis was to determine the extent to which they were equivalent. It should be recalled that the five sets of tests were administered to five groups of students on random basis in the class. That is, each of the test was taken by a section of students in each class randomly divided into five. The five groups of the students were equivalent groups. It should also be recalled that Kulkahi (1969) reports that for the forms of the tests to be equivalent the means

and standard deviations obtained by the groups would be about the same. To this effect the means, standard deviations and probability coefficients of the five sets of tests were calculated using t - test. The results showed that sets 1 and 5; and sets 2, 3 and 4 were equivalent by virtue of their means and standard deviations as indicated by the Table 4.3 below. . Also the results showed a high correlation at 0.01 level of significance for all the five sets of tests. This implies that all the five sets of tests can be said to be parallel in that there is no significant difference between the means and the standard deviations of all of them. Hence a reliable samples of the course content under study. It also further supported the content validity of the tests and indicated that the course objectives are achievable.

Table 4.3 Significance of Means and Variance

				N
1.	$\bar{X}_1$=23.96	t=.77 Prob=.443	σ_1=23.40	F=1.17 Prob=.382	94
	$\bar{X}_2$=21.69		σ_2=21.59		147
2.	$\bar{X}_1$=23.96	t=.80 Prob=.370	σ_1=23.40	F=1.14 Prob=.471	94
	$\bar{X}_3$=21.30		σ_3=21.91		151
3.	$\bar{X}_1$=23.96	t=.57 Prob=505	σ_1=23.40	F=1.24 Prob=.233	94
	$\bar{X}_4$=21.99		σ_4=21.03		139
4.	$\bar{X}_1$=23.96	t=.13 Prob=.899	σ_1=23.40	F=1.08 Prob=.697	94
	$\bar{X}_5$=23.56		σ_5=22.56		128
5.	$\bar{X}_2$=21.69	t=.15 Prob=.877	σ_2=21.59	F=1.03 Prob=.697	147
	$\bar{X}_3$=21.30		σ_3=21.91		151
6.	$\bar{X}_2$=21.69	t=.12 Prob=.906	σ_2=21.59	F=1.05 Prob=.755	147
	$\bar{X}_4$=21.99		σ_4=21.03		139
7.	$\bar{X}_2$=21.69	t=.27 Prob=.786	σ_2=21.59	F=1.09 Prob=.607	147
	$\bar{X}_5$=23.56		σ_5=22.56		128
8.	$\bar{X}_3$=21.30	t=.27 Prob=.786	σ_3=21.91	F=1.09 Prob=.624	151
	$\bar{X}_4$=21.99		σ_4=21.03		139
9.	$\bar{X}_3$=21.30	t=.85 Prb=.398	σ_3=21.91	F=1.06 Prob=.730	151
	$\bar{X}_5$=23.56		σ_5=21.56		128
	$\bar{X}_4$=21.99	t=.59 Prob=.557	σ_4=21.03	F=1.15 Prob=.419	139
	$\bar{X}_5$=23.56		σ_5=22.56		128

Meams and Standard Deviations not significant at .01 level of significance.

SHIM IBRAHIM LIB

The second analysis was to determined the difficulty level and discriminating power. As the marks of the 27 items range from 2 to 8 a candidates was considered to have got an item correct if he/she got the 2 marks in all the items carrying a maximum of 2 marks; 2 marks in all items carrying maximum of 3 marks; 3 marks in all items carrying a maximum of 4 marks; 4 marks in all the items carrying a maximum of 5 and 6 marks; and 6 marks in the item carrying a maximum of 8 marks for the purpose of computing the item difficulty level and discriminating power.

The item difficulty level was calculated for each set of test using all the candidates who took each test. It should be recalled that the index of items difficulty indicated the percentage of candidates who scored an item correctly. For this study the following statistical formula was used to determine the difficulty level at the second try-out.

$$P = \frac{R}{T} \times 100 \text{ where}$$

P = difficulty index

R = the number of candidates who got the item right

T = the total number of candidate who tried the item.

Gronlund (1981;258).

The discriminating power of an item differentiates between students with high and low achievement. It is obtained by subtracting the number of students who got an item correctly in the lower group from the number of students who got the item correctly in the upper group and dividing by the number of students in each of the two groups. The determination of the upper group and lower group is usually done by selecting some random sample of the students who took the test especially where the sample is large; Brown (1976). For example if 275 candidate took a test, 100 candidates may be randomly selected for the calculation of the discriminating index. On the other hand the total number of candidates who took the test can be used for classifying students into upper and lower groups for the purpose of calculating the discrimination index for each item. When this is done, it is known as the internal consistancy method of computing indices of item discriminating power; Ahmann & Glock (1963). A decision should then be made as to which part of the distribution of scores is the upper group and which part is the lower group. Some select the lower and upper halves of the distribution, others the lower and upper third, discarding the middle third while others select the lower and upper 27% discarding the middle 46%. Ahmann and Glock however report that the last selection has

received a great deal of support by testing specialists.

For the purpose of this study the discriminating index was calculated using all the candidates who took each set of the test with lower and upper 27% of the distribution. The following formula was used:-

$$D = \frac{U-L}{N} \text{ where}$$

D = discrimination index

U = number of students in the upper group who scored the item right

L = number of students in the lower group who scored the item right

N = number of students in each of the two groups.

Generally for standardized achievement test, items with difficulties of .5 are considered optimal. But there are cases when one would use warm up items at the beginning of a test to build confidence and get the students to the testing condition. Such items should be quite easy. On the other hand if the tests are designed to discriminate at a particular point the average item difficulty should be close to that point. Thus if one wanted to differentiate the top 20 percent of a group from the rest of the students, the best approach would be to use items that only 20

percent of the candidates answer correctly; that is items with an average difficulty index of around .2; Brown (1983). Similarly for the discriminating power, an index of .4 and above is considered to be ideal, but for practical consideration as a rule of thumb we can say that any indices between .15 to .20 or higher would indicate good discrimination; Brown (1983).

This study, therefore considered the difficulty level of .09 and above for the final tests. This will allow the test user to differentiate at around 10% of the group who would likely be placed for special classes that might require more mathematics. Also the discriminating index of .25 and above were considered for the final tests. The tables for difficulty level and discrimination index are in Appendix E

4.6 The Final Test.

The results of the item analysis helped the researcher to select the items that were finally administered to 3081 students of 14 institutions all over the state as indicated in Table 3.3. Two sets of parallel tests were prepared. As in the second try-out, each test consisted of 27 short-answer and essay type items of varying complexity in line with the guideline

already developed earlier in this chapter. The tests were administered and scored in various institutions by the teachers using the directions and marking schemes that were collected by them from the researcher. The Administration of the tests took place at different dates in the various institutions, between June and July, 1983. This is because majority of the schools which took the tests wanted to use the results of the tests for their end of term examinations which were time-tabled for different dates in different institutions. All teachers involved in marking the tests were properly co-ordinated as to how and where to award marks in the students solutions.

All the test papers and results sheets were returned to the researcher except that of Rumfa College in which only the results sheets were returned. The reason was that the Principal directed the teachers to give the marked papers back to the students and they were never retrieved. Unfortunately again, the results sheets of the same college were parts of the properties stolen when the researcher was attacked by armed robbers in his Sharada Residence before the information was stored into the computer. Hence the researcher's inability to use the results of fourteen institutions as shown in Table 3.3 for the final analysis.

4.7 Statistical Analysis of Final Testing.

The analysis of the test results involved two main statistical tests.

i. the determination of Norms; and

ii. reliability estimate.

(i) Norms: It should be recalled that one of the main objectives of this study was to establish norms. The test scores are interpreted through the establishment of:

(a) Class norms for all the distribution,

(b) Sex norms, and

(c) Institution type norms.

The scores were then expressed in terms of frequency distributions by norms with tables and graphical representations. Also percentile rank, normalized standard scores, skewness and kurtosis as well as standard errors of measurement of each norm were represented. ABU Computing Facilities were used for the analysis of the results.

Frequency Table. The scores obtained by each of the whole group and of the sub-groups in each test are arranged in successive range on the table. This shows how many candidates fall within each range.

Graphical Representation. The frequency polygon of the distribution of each test scores for the whole group and sub-groups are presented. This is to determine the normalcy of the distribution for the whole group and sub-groups in each test.

Percentile Ranks. Percentile Ranks for the whole group and for each sub-group are calculated by using the method suggested by Brown (1983); That is

(a) Preparing a frequency distribution of the scores

(b) Finding the cumulative frequency (CF) of the lower limit of each score. This is done by finding the sum of all scores below the scores in question.

(c) Finding the cumulative frequency to the mid-point of the score interval (CFmp). This is obtained by adding one-half the number of scores in the interval to the CF.

(d) Finding the cumulative proportion (CP) by dividing CFmp by n, the total number of scores.

(e) In finding the percentile rank (PR), we multiply by 100.

The above steps are put in tabular form below:

Table 4.4 Computation of Percentile Rank.

Scores	F	CF	CF_{MP}	CP	PR
X_1	f_1	CF_1	CF_{mp_1}	CP_1	PR_1
X_2	f_1	CF_2	CF_{mp_2}	CP_2	PR_2
⋮	⋮	⋮	⋮	⋮	⋮
X_n	f_i	CF_j	CF_{mp_k}	CP_n	PR_m

Normalized Standard Scores. Since the raw scores of this study are not normally distributed the researcher used normalized standard score procedures in order to adjust the scores into normal distribution as advised by Brown (1976). To do this the following steps were employed.

(a) From the computation of percentile rank find the commulative proportion (CP) of the raw score.

(b) Using the Table of the Areas of the Normal curve, we find the Z-score comparable to the CP. For this study the table prepared by Glass and Stanley (1970) is used.

(c) We then transform the scores to another scale. We use area transformation with a mean of 50 and a standard deviation of 10 points. That is $T = 50 + 10Z$.

The Table below shows the tabulating procedure.

Table 4.5 Computation of Normalized Standard Scores.

Raw Scores	Commulative Proportion (CP)	Standard Score (Z)	Transformed St.Sc. (T).
X_1	CP_1	Z_1	T_1
X_2	CP_2	Z_2	T_2
⋮	⋮	⋮	⋮
X_n	CP_i	Z_j	T_1

Skewness and Kurtosis. The rational for calculating the Skewness of the scores is to determine the nature of the distribution against the normal curve. It is calculated using the formula.

$$g = \frac{\text{mean-median}}{\text{Standard Deviation.}}$$

When the distribution is symmetrical mean minus median is zero, but when the distribution is asymmetrical

the g. $\neq$ 0. Also when the distribution is positively skewed g is positive, when negatively skewed g is negative.

The concept of Kurtosis is closely linked to the relative thickness of the tail of the distribution. It is calculated to determine the normalcy of the curve. For a normal distribution the Kurtosis is 0. For a leptokurtic distribution the Kurtosis is greater than zero and for a platykurtic distribution the Kurtosis is less than zero.

(ii) Reliability Estimate. This is calculated using Pearson Product-Moment based on split-half. The researcher was unable to administer the two sets of tests to the same group of students during the data collecting process because of the large number of students involved in the study. This have allowed the correlation for equivalence to determine the extent to which the two tests are parallel by Product-Moment only. The researcher is aware of the fact that few students in a study like this can be administered the two sets for the purpose of determining the estimate for equivalence. But he is again unable to do this.

Pearson Product-Moment. Each of the two sets of tests has been split into two halves; odd-numbered and even-

numbered items by norms. Since there are 27 items in each test, the two halves contain 13 items each by setting the 27th item to blank. Product Moment formula as stated below is applied to obtain the correlation of the half-length test value of each test by norm.

$$r = \frac{N\Sigma xy - \Sigma x\Sigma y}{\sqrt{[N\Sigma x^2-(\Sigma x)^2][N\Sigma y^2-(\Sigma y)^2]}}$$

Spearman Brown. In order to get the total test reliability coefficient, the reliability estimate obtained by spliting each test by the use of Product-Moment is stepped up into full-length value. This is done by the use of Spearman Brown formula:

$$r_2 = \frac{2r_s}{1+r_s}$$

The estimates obtained by both Product-Moment and Spearman Brown are presented in chapter 5 by test set and norms.

Standard Errors of Measurement. The reliability of the tests was further investigated by examining the standard error of measurement of the tests for each group. Lord (1955) defines the standard error of measurement associated with Psychological testing as:-

> ...the standard deviation of scores an individual might be expected to obtain on a large number of parallel test forms.
>
> Ferguson (1981; 443).

It should be noted that one of the main objectives of psychological testing is to have the observed scores reflect the true scores with as little error as possible. In order to get to the true scores the obtained scores should approximate to, as near as possible, the true scores of the individual's performances. The assumption here is that the ability of the individual remains unchanged and is not affected by variables such as practice, maturity, ambiguity of the items and the like. In other words by calculating the standard error of measurement we are trying to determine the reliability estimate of the tests.

Chapter 5

DATA PRESENTATION AND ANALYSIS.

5.1 Introduction.

This chapter is concerned with the organization, statistical analysis and interpretation of the data obtained from this study. It establishes the various norms as discussed in the previous chapters. These include the class norms, the sex norms(Boys and Girls) and the school type norms (Private and State Government.) Percentile Ranks and Normalised Standard Scores are established from the study data obtained from two sets of tests on algebra administered to class 3 students in Kano State. Altogether five norms have been established for each test set.

5.2 Frequency Distribution of the Study Data.

i. Means and Standard Deviations: The students' mean scores and standard deviations on each test set and for each group is presented in Table 5.1 below. It should be recalled that one method of establishing the degree of equivalence of two tests is to observe that the means and standard deviations of the test scores to be about the same, Brown (1976) and Kulkarni (1968). From this study, therefore, the

performance of two groups norms out of the five norms established in this study have indicated that the two forms of the tests are equivalent. These are the performances of Girls Institutions and the Private Institutions. The remaining norms have somehow shown low degree of equivalence which may be attributed to the kind of attention the students received from either the school administrations or the teachers. It was observed during the visits to the schools and the workshop that most Government Girls Institutions had more qualified mathematics teachers than the boys.

Table 5.1 Means and Standard Deviations of The Various Norms and Test Set.

Test Set		Total	Boys *	Girls *	Private	State Government
	$\bar{X}$	7.94	6.36	7.65	13.93	5.27
A	SD	9.86	9.18	9.28	11.87	7.40
	N	1356	838	319	418	938
	$\bar{X}$	6.51	4.36	7.16	13.91	4.15
B	SD	9.38	7.71	9.9	11.66	7.06
	N	1300	778	356	314	986

*Mixed Institution not included in the sex norms.

The mean scores of Private norms is higher than any of the other type of norms. This could also be attributed to the fact that students in these type of schools have better teaching facilities and

better teachers especially that the schools recruit and maintain their own teachers. The relatively higher mean scores of girls norms in relation to the other norms apart from the Private norms can be accounted by the observation made earlier. That is having more qualified teachers.

Table 5.2 Significance Of The Difference Between Means.

Test Set	Variables	$\bar{X}$	SD	T.Value	DF
A	Boys	6.36	9.18	2.115^{x}	1155
	Girls	7.65	9.28		
	Private	13.93	11.87	13.746^{xx}	1354
	State-Govt.	5.27	7.40		
B	Boys	4.36	7.71	4.667^{xx}	1132
	Girls	7.16	9.90		
	Private	13.91	11.66	14.145^{xx}	1298
	State-Govt.	4.15	7.06		

x Significant at .05 level of Significance

xxSignificant at .001 level of Significance.

The tests of the significance of the difference between two means in Table 5.2 indicate that there is a significant difference between girls' and boys' achievements in algebra at .05 level of significance for Set A and at .001 level of significance for Set B. Also significant difference exists in algebra achievements between students in Private Schools and

State Government Schools at .001 level of significance for Set A and Set B tests. This helps to justify the establishment of the different norms.

ii) Skewness And Kurtosis:- The Skewness and Kurtosis of the distribution is presented in Table 5.3 The Table indicates that the distributions of the various norms for both sets of tests are positively skewed. This shows that the tests are difficult ones and most students scored very low. It has already been indicated that the tests have been prepared to help determine the top 10% of the students population who should be placed in the appropriate type of Senior Secondary Schools that may demand more mathematics than others, such as Senior Secondary Schools Science or Technical.

Table 5.3 Skewness And Kurtosis of Scores Distribution Of The Various Norms And Test Set.

Test Set		Total	Boys	Girls	Private	State Govt.
A	Skewness	1.98	2.71	1.81	1.42	2.28
	Kurtosis	5.33	10.41	4.03	3.06	6.47
B	Skewness	4.22	2.96	2.61	1.37	2.87
	Kurtosis	19.11	10.38	10.66	3.14	10.46

The Curves of the distributions are, therefore, asymmetrical as observe from the Table. The values of the Kurtosis indicate that the distributions, are leptokurtics since all the values are positive.

(iii) Frequency Tables And Frequency Polygon:- The raw scores of the distributions have been tabulated using the

class-interval of five points. The tabulations have been done for all the five norms identified above. This gives ten tables, two for each norms based on the two tests as shown in Appendix H. To give more meaning to the frequency tables, a frequency polygon is presented for each test set to show the nature of the distribution for each norm. The frequency polygons are presented in Figs. 5.1 - 5.10.

On the whole the distribution of scores shows that the two sets of tests are very difficult for the population under study. In all the norms, the students' performances are very low. Most students in the various norms scored less than 5%. In fact in all the norms but one, that is the Private Institutions over 50% of the students scores less than 5%. It can also be observed that less than 11% of the students who took these tests scored upto 20% in the tests in the four norms. While in one norms which is the Private Institutions about 24% of the students scored upto 20% in the tests. This makes the tests suitable and acceptable for determining the top 10% of the students population for placement to Senior Secondary Schools that may require more mathematics.

FIG. 5.1

FREQUENCY POLYGON ON SET A.
TOTAL.

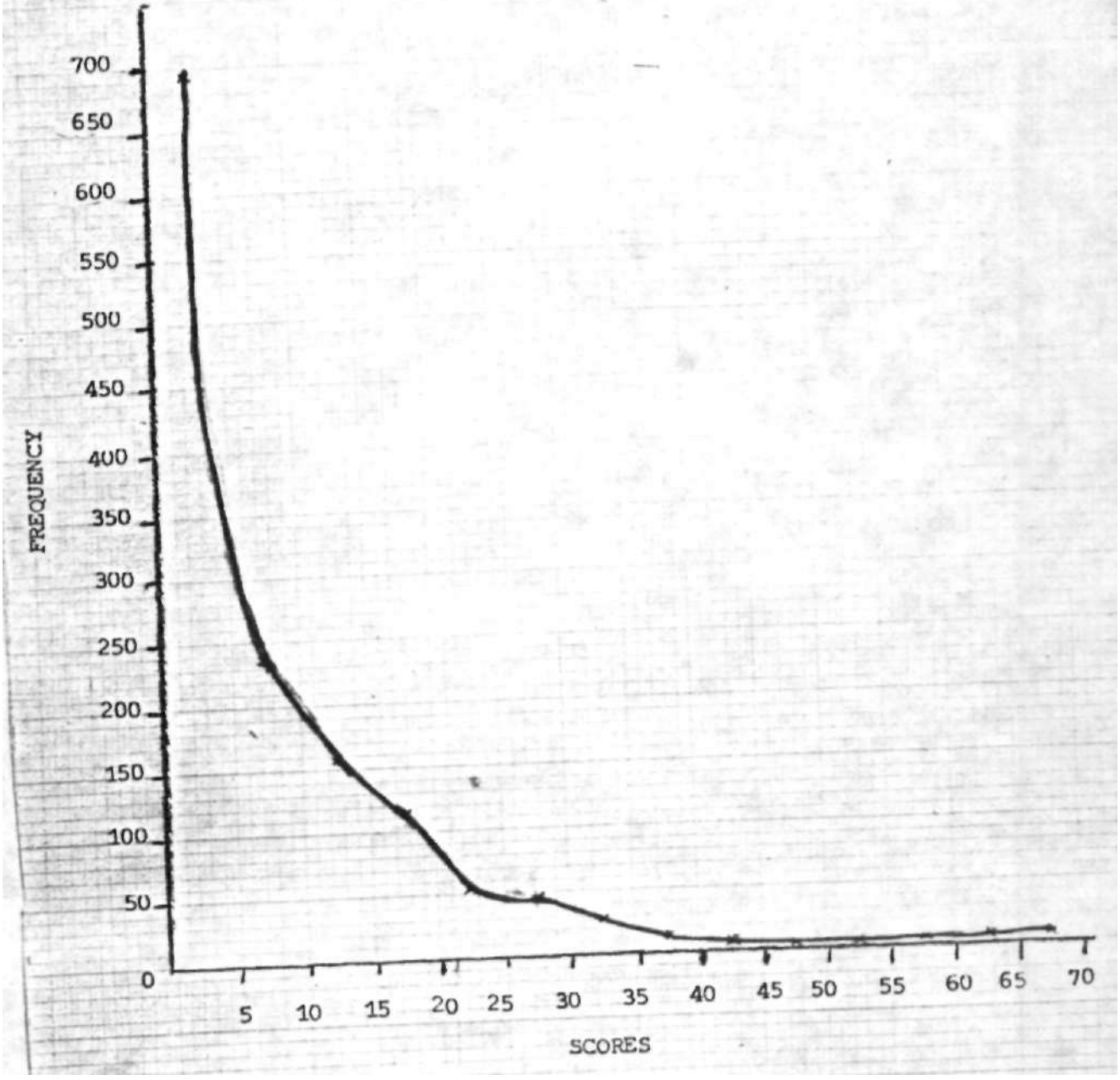


FIG. 5.2

FREQUENCY POLYGON OF SET B
TOTAL.

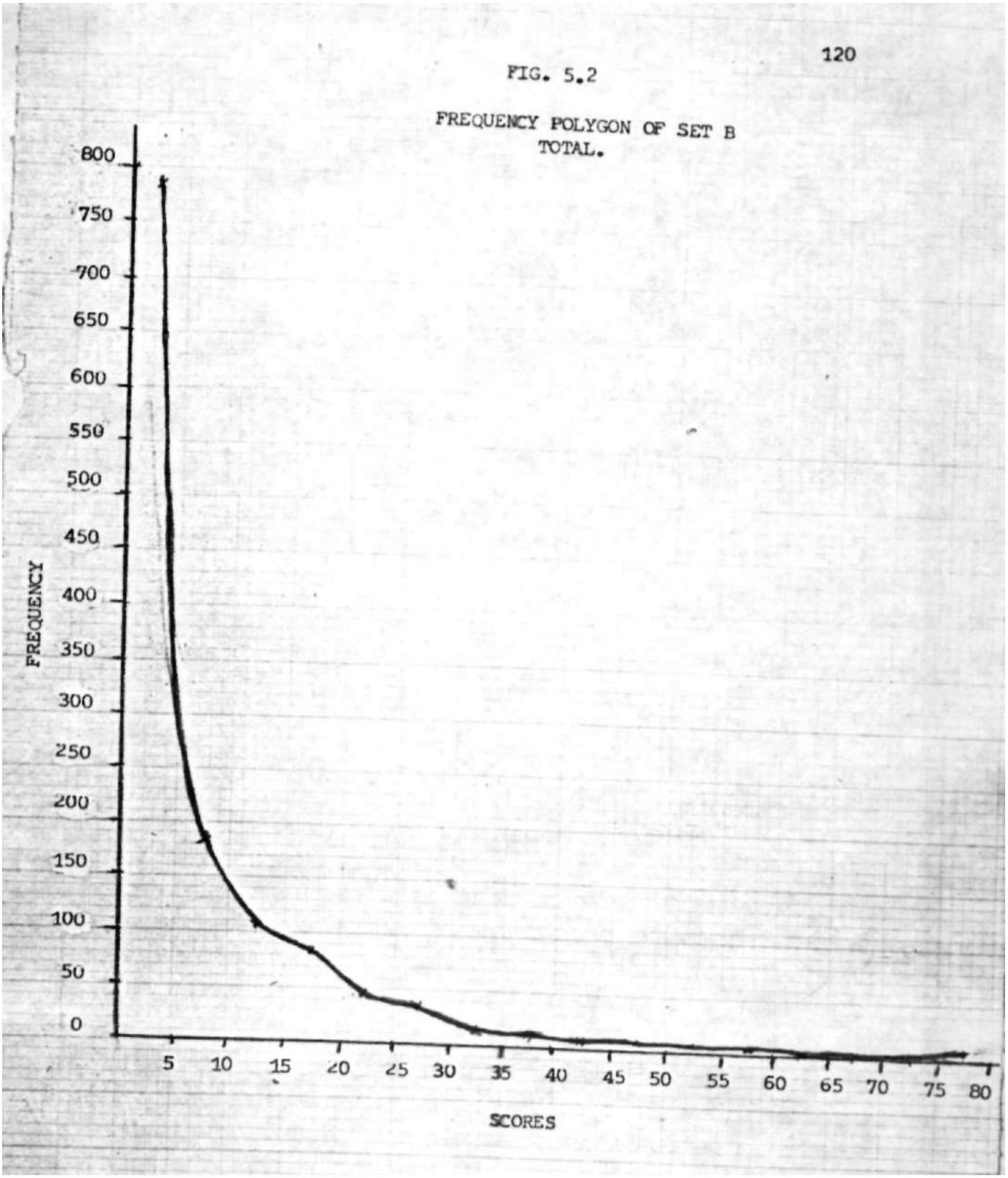


FIG. 5.3

FREQUENCY POLYGON OF SET A
BOYS

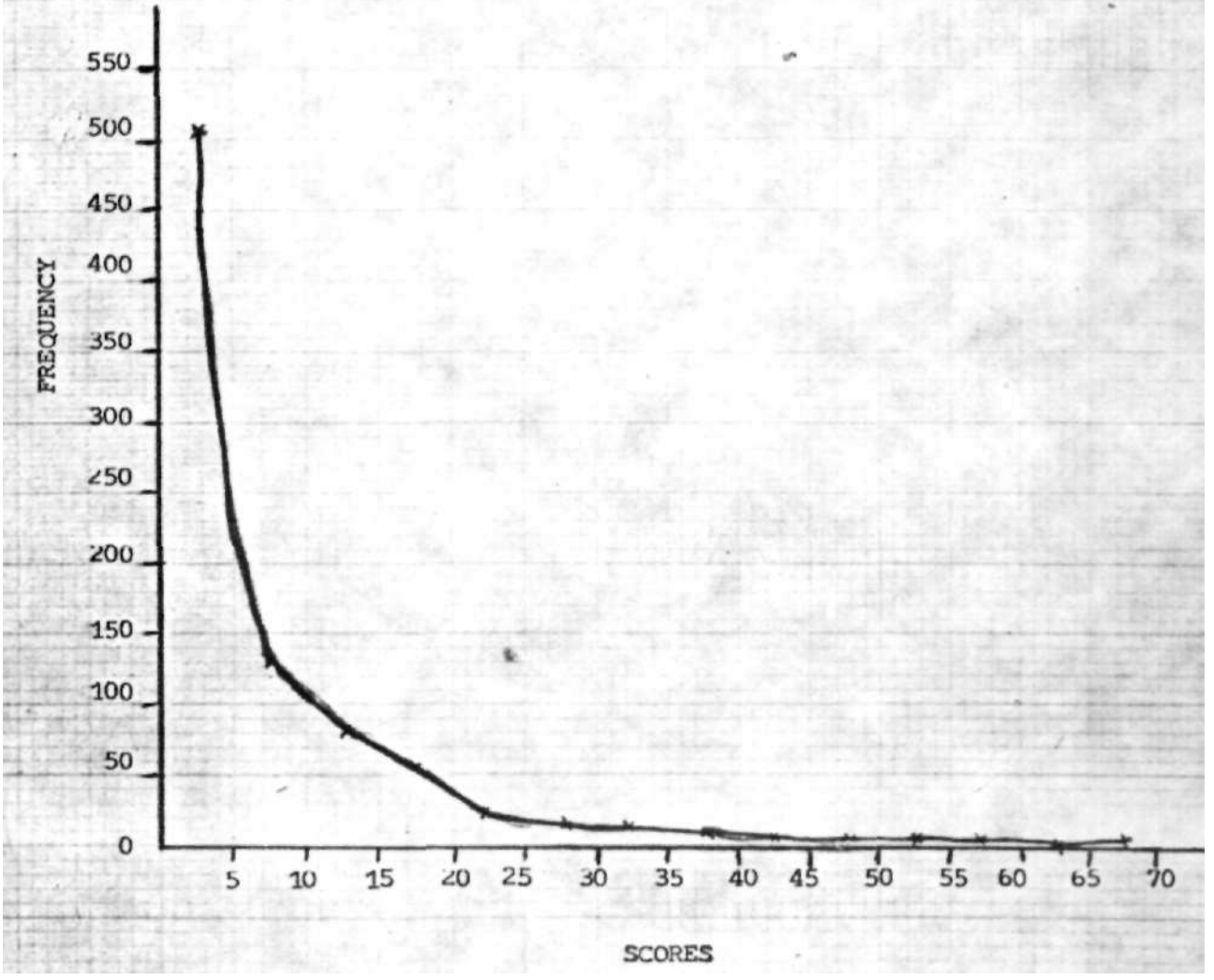


FIG. 5.4

FREQUENCY POLYGON OF SET B
BOYS

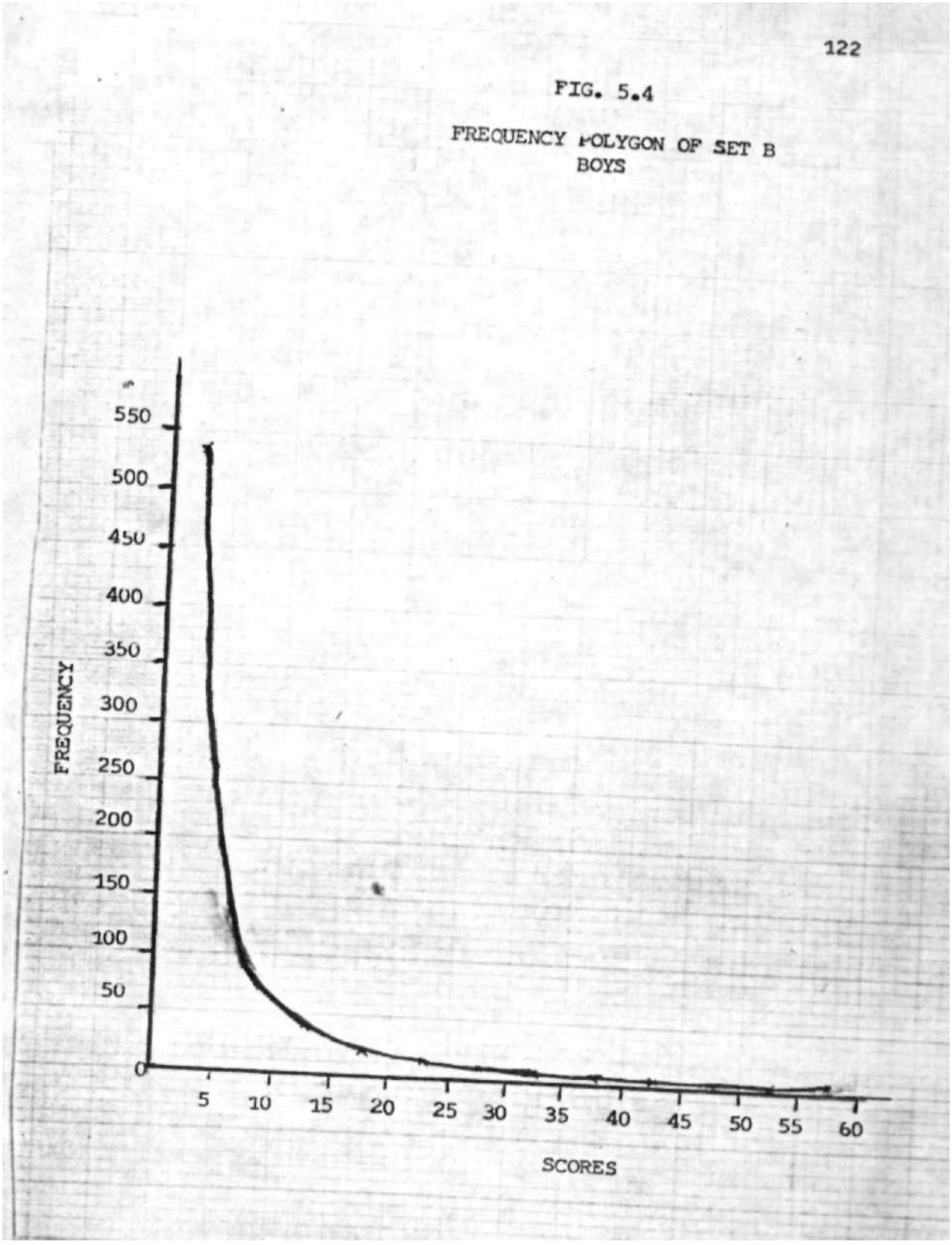


FIG. 5.5

FREQUENCY POLYGON OF SET A.
GIRLS

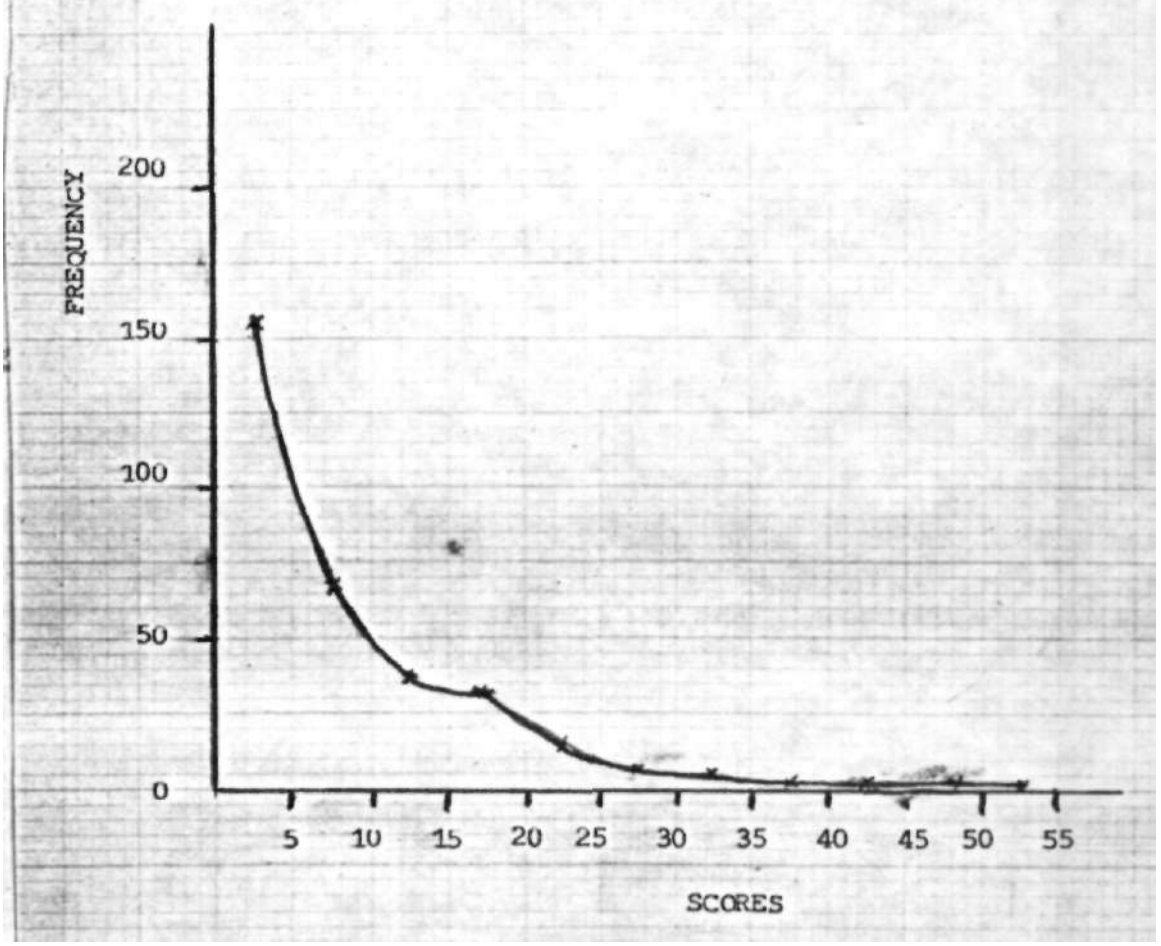


FIG. 5.6

FREQUENCY POLYGON OF SET B.
GIRLS

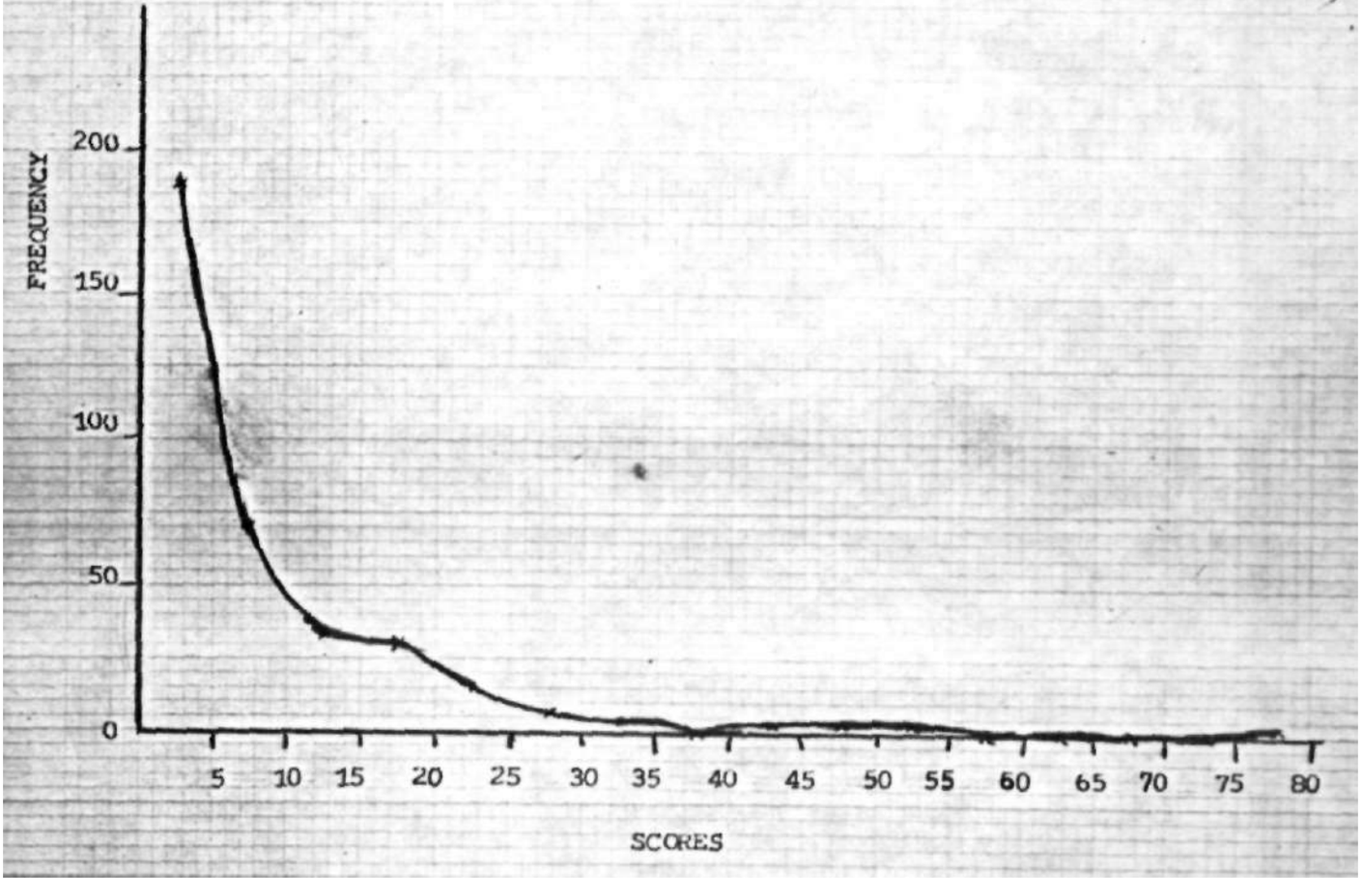


FIG. 5.7

FREQUENCY POLYGON OF SET A.
PRIVATE

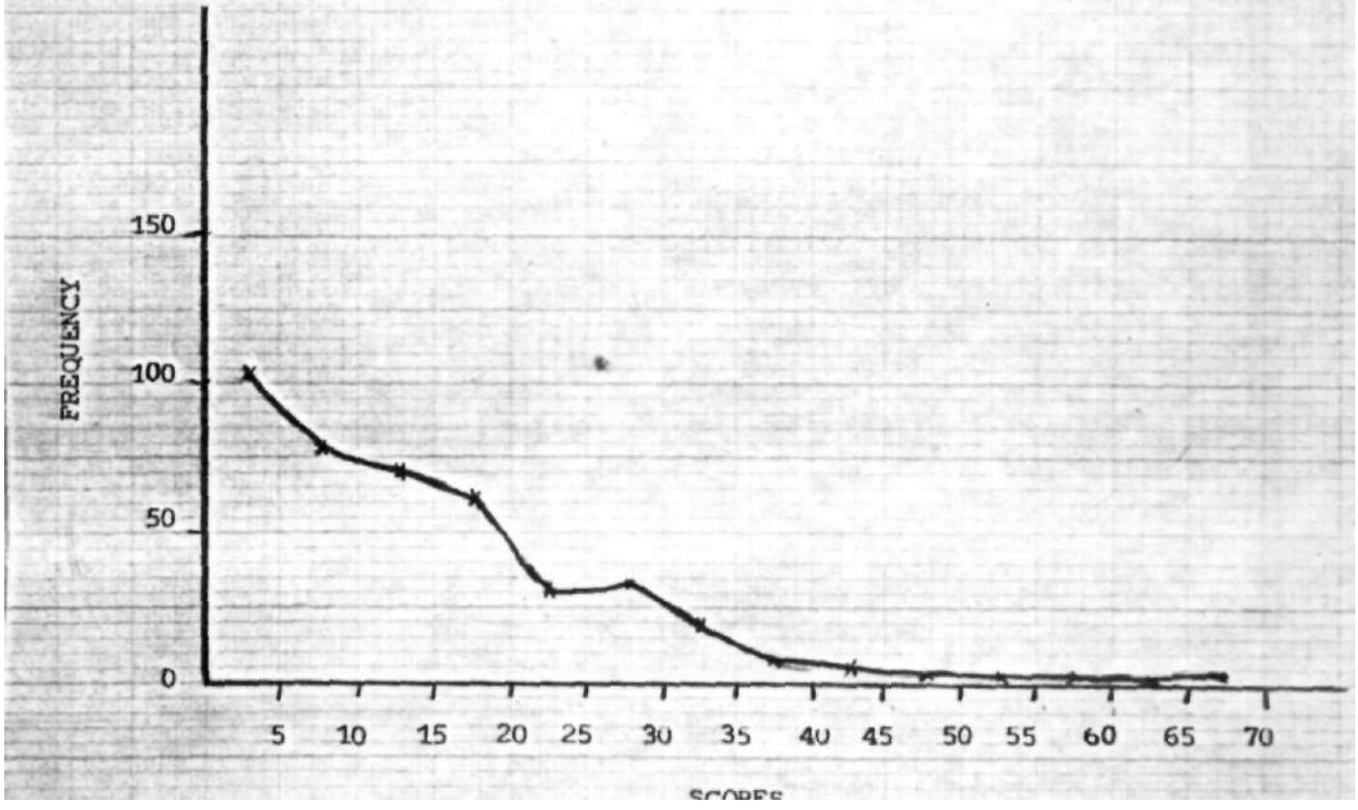


FIG. 5.8

FREQUENCY POLYGON OF SET B.
PRIVATE

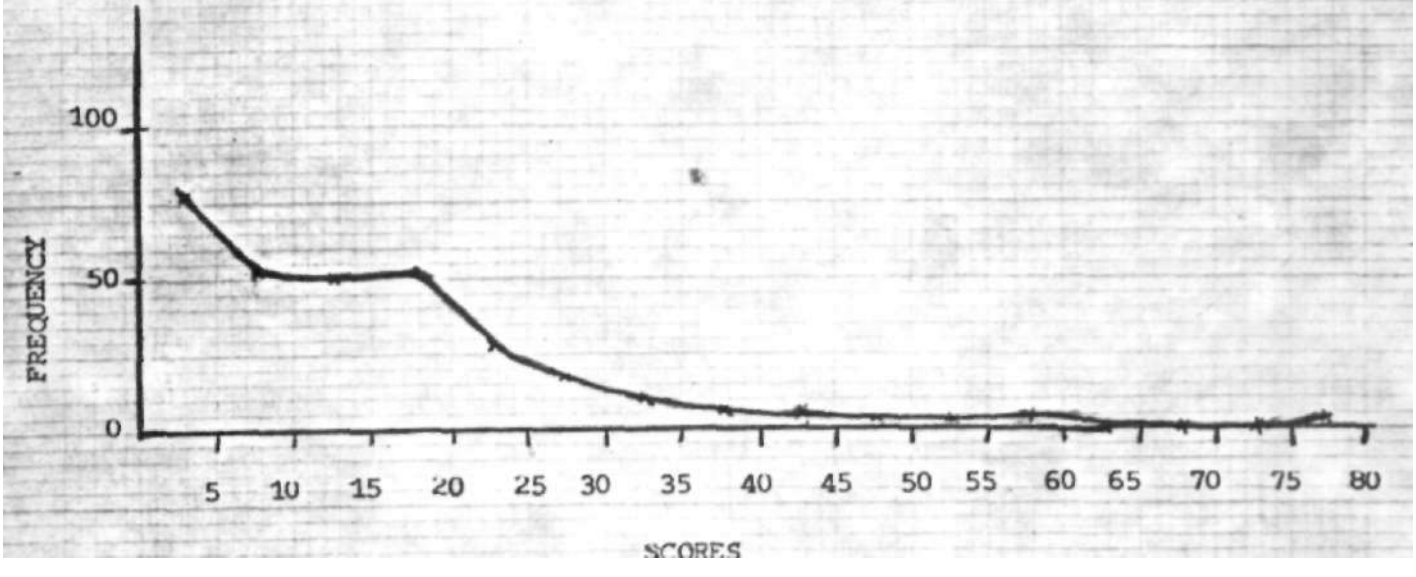


FIG. 5.9

FREQUENCY POLYGON OF SET A.
STATE GOVERNMENT.

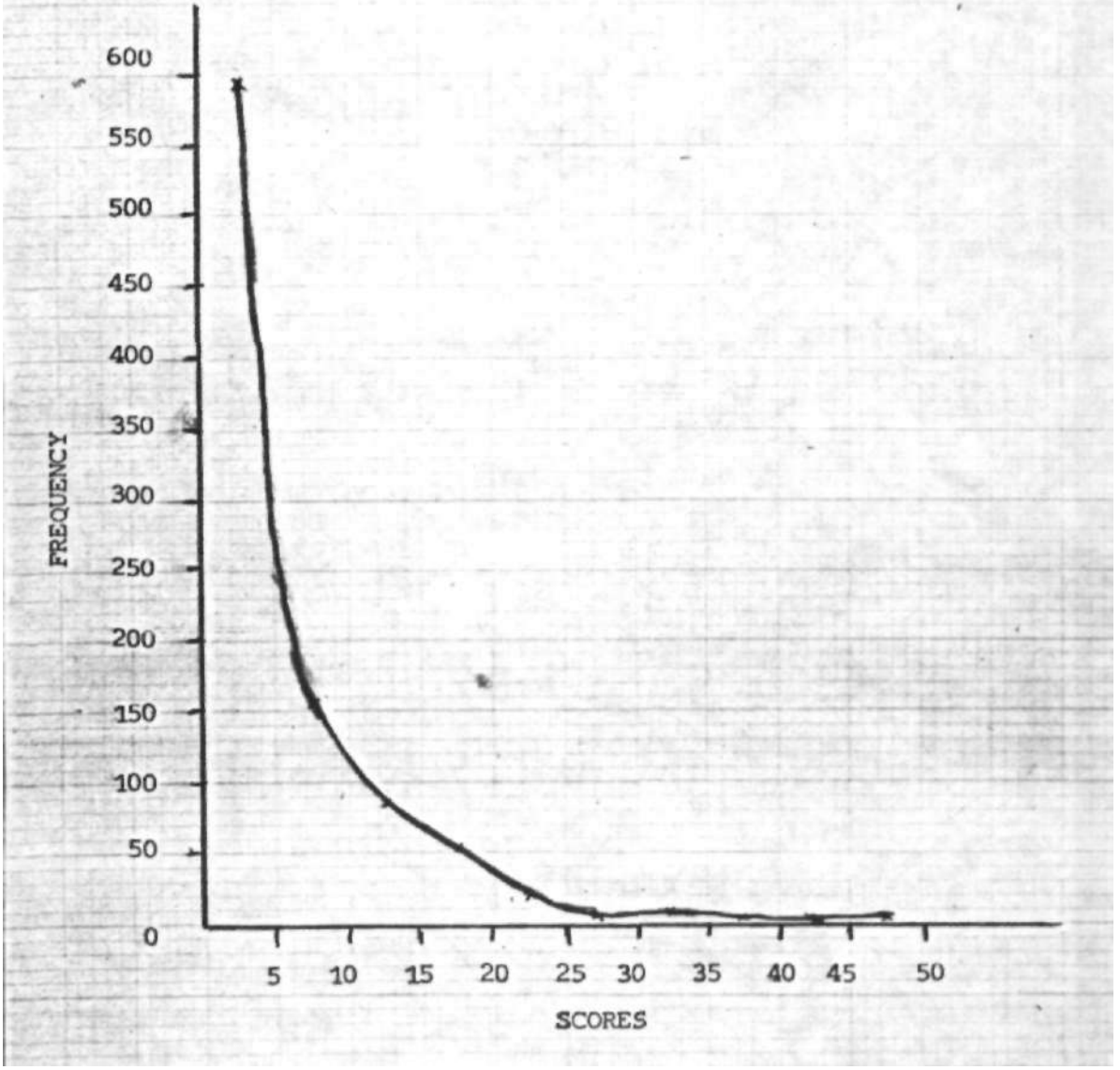
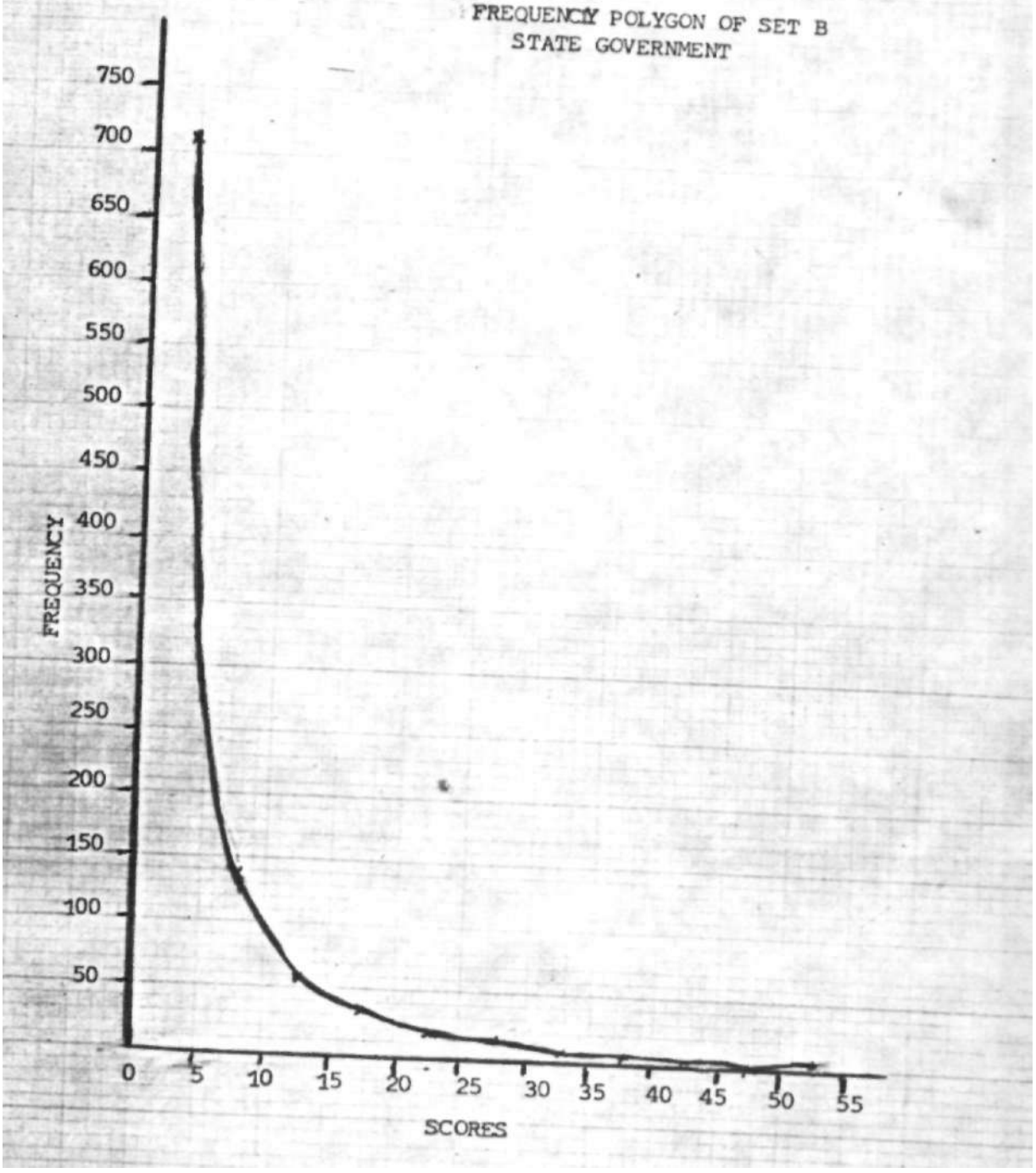


FIG. 5.10

FREQUENCY POLYGON OF SET B
STATE GOVERNMENT



5.3 Percentile Ranks Of The Test Scores.

The percentile ranks of the test scores have been calculated for each norms in the two sets of tests. They are presented in Appendix I. The calculation followed the procedure described in the last chapter. The percentile ranks of all the norms are generally high. This might be attributed to the fact that majority of the students who took the tests scored very low and the distribution of the tests scores cluster around the low percentile points which are represented by very high frequencies.

The tables present the transformation of the observed scores to the percentile ranks. The students raw scores can be interpreted to the percentile ranks and vice-versa. For example, in the percentile rank table of set A for total the raw scores of 20-24 are represented by the percentile rank of 90.5. That is any student scoring marks between 20-24 performs higher than 90.5% of class 3 students in Kano State and Lower than 9.5% of the class 3 students in set A. While the same scores in the same set of test in Private Institutions show the percentile rank of 78.5. Which means that the student in Private Institution only performs higher than 78.5% of the class 3 students in all the Private Institutions in the State.

5.4. Normalized Standard Scores Of The Distribution.

The raw scores of the students performances have again been transformed into normalized standard scores. The calculation procedures employed have been described in the last chapter.

Appendix J shows the convertion of raw scores to normalized standard scores for each norm and test set. The use of normalized standard scores here is only for comparing the raw scores with the T- Scores. Otherwise transforming scores into Z- Scores should only be done with the scores that are normally distributed. Since the distribution in any of the norms is nowhere near normality, percentile ranks are more appropriate and should be used for interpretating and using results in this study.

5.5. Reliability Estimate.

The students' scores on the odd-numbered and even-numbered items are correlated in each of the two tests by norms using the Pearson Product-Moment formula:-

$$r = \frac{N\Sigma AB - \Sigma A\Sigma B}{\sqrt{\left[N\Sigma A^2-(\Sigma A)^2\right]\left[N\Sigma B^2-(\Sigma B)^2\right]}}$$

The half-length test reliability Coefficients are shown in Table 5.4

Table 5.4 Half-Length Reliability Coefficients of Each Test By Norms.

Test Set		Total	Boys	Girls	Private	State Government
A	r	.7187	.7580	.6400	.7092	.6459
	N	1356	838	319	418	938
B	r	.7758	.7955	.7349	.7487	.7240
	N	1300	778	356	314	986

All the reliability coefficients are significant at .001.

The half-length reliability coefficients are then corrected with the Spearman Brown formula in order to obtain an unbiased estimate of the reliability of the full-length test.

$$r_2 = \frac{2r_s}{1+r_s}$$

Where r_2 is the reliability coefficient of the full-length tests;

r_s is the reliability of the half-length test value.

The full-length test reliability coefficients are shown in Table 5.4

Table 5.4. Reliability Coefficients of Each Test By Norms.

Test Set		Total	Boys	Girls	Private	State Government
A	r_2	.84	.86	.78	.83	.78
	N	1356	838	319	419	938
B	r_2	.87	.89	.85	.86	.84
	N	1300	778	356	314	986

It is noteworthy that eight of the ten correlation coefficients are all within the .80's and two are in the late .70's. It can therefore be claimed that the tests are good. Anastasi (1976), Adam (1966), Brown (1983) and Gronlund (1981) suggest that reliability coefficients of good tests should usually fall in the .80's or .90's. This is moreso with standardized tests.

It can be observed from the Table that the reliability coefficients of set B test is consistently high in all the norms. While relatively low reliability coefficients in set A test for Girls and State Government Schools are realized. It can, therefore, be inferred that Set B test is a more consistant test than the Set A test for all the norms, even though the differences are not significant. Also the reliability coefficients obtained for total (Grade), Boys and Private norms in both two tests are consistantly high, especially that of Boys' norm. This implies that the two tests can be administered inter-changeably with a very high degree of confidence. Even in the norms where comparatively low coefficients are realized in Set A test, the virtue of the equality of the means and standard deviations as obtained in Table 5.1, especially for Girls, make the inter-change possible with high degree of confidence.

On the whole, it can be concluded that the tests are consistant, inter-changeable and parallel.

5.6. Standard Error of Measurement

The Table below shows the standard error of measurement SE_m for each norms.

Table 5.6. Standard Error of Measurement of Each Test By Norms.

Test Set		Total	Boys	Girls	Private	State Government
A	SE_m	0.27	0.32	0.52	0.58	0.24
	N	1356	838	319	418	938
	SE_m	0.26	0.28	0.52	0.66	0.23
	N	1300	778	356	314	986

It can be observed from the Table that, the margin of error surrounding the individual's test score is very low indeed for all the norms. One is therefore safe to claim that the tests do infact measure the true individual performances. Thus, strengthening the reliability of the tests for all the norms.

Chapter 6.

DISCUSSION, MANUAL INFORMATION AND RECOMMENDATIONS.

6.1 Discussion.

The Data obtained from this study have been presented, analysed and interpreted in chapter 5. These show the practical procedure for standardization process. This chapter provides some further discussion of the results and give some guidelines for using the tests to whoever wants to use them for the purposes they are developed.

(i) Distribution of the Study Data. The means and standard deviations of the distribution by norms indicate different degree of equivalence. Two of the norms, that is, Private and Girls, show very high degree of equivalence by virtue of their means and standard deviations as suggested by Brown (1976) and Kulkarni (1968). The Girls' norm has .49 and .62 as the difference between the means and standard deviations of set A test and set B test respectively. While the Private has .02 and .21 accordingly. The difference in means and standard deviations of set A test and set B test in Grade (Total) norm is 1.43 and .48 respectively while State Government norm has 1.12 and .34 accordingly. These show that by virtue of standard deviation in the two norms the

two tests are very close to being highly parallel. The Boys' norms, however, shows wider range between the means and the standard deviations of the two tests; 1.12 and .34 have been observed respectively.

It has, again been observed that in all the norms but Private the standard deviations are greater than the means. This implies that the distribution of scores concentrate around zero with very wide range. The results of this study also show that girls have higher achievement in algebra than boys, all factors contributing to the success of teaching and learning activities not withstanding; and Private Institutions perform better than State Government Institutions.

(ii) Percentile Ranks of The Study Data:- The scores of the students performances in the two tests have been converted into percentile ranks by norms. Ten tables have been presented in chapter 5 to show the conversion of the students raw scores to the corresponding percentile rank. (Tables 5.3-5.12). It is observed that, the students' performances in both two tests have been very low indeed. Majority of the students in all the norms scored less than 5%. This accounts for sudden drop of the cumulative proportion value at the class-interval of 0-4 in each norm. This is not surprising when we consider the students'

performances in GCE/SC and NTI examinations. Year in and year out we learn and talk of students' poor achievement, especially in Mathematics, thus indicating to some extent the validity of the tests.

(iii) Normalized Standard Scores Of The Study Data:

Normalized standard scores are used to force raw scores into normal distribution when scores are not normally distributed. But this is done when the distribution approaches normality as suggestd by Brown (1976). In this study, the distributions in all the norms are not in any way near normality. It should therefore, be noted that in interpreting the results of this study, normalized standard scores presented in the study should not be used for fear of distortion of the results. Percentile ranks should be used instead.

(iv) Reliability Estimate: The reliability coefficients of the two sets of tests were calculated using split-half based on Pearson Product Moment formula and corrected by Spearman Brown formula for each norm. It has been found that 8 of the 10 coefficients are in the .80's while only 2 are in late .70's. It is observed that the two sets of tests are reliable and one can be a substitute of the other.

The reliability estimate was further ascertained by considering the standard error of measurement of each test by norm. The errors surrounding the performances of the students in the tests by norm range from .23 - .66. It is observed that the individual scores can be described as producing consistant measures. Hence further strengthening the reliability of the tests.

6.2. Manual Information

The standardization process of any test will be complete only after a test manual has been prepared. Such a manual contains the technical process of the test construction. Brown (1983) explains that a test manual includes the definition of the domain, test blue-print or table of specification, basis for item selection and item statistics as well as normative data, reliability and validity data. Some detailed directions for administering and scoring of the test, as well as some interpretive materials and use of the test results should also be included in the manual to help the test user makes proper and good use of the test.

For this study the test manual has been developed in chapters 4 and 5. It is further explained as follows:-

(a) The Test Domain:- The domain of the Secondary Three Algebra Test (STAT) of this study consists of all the topics under the Algebraic Processes of the contemporary mathematics curriculum for Junior Secondary Three. The main sub-topics in the domain are:

i. Factorization;

ii. Solution of Fractional Equations;

iii. Solution of Simultaneous Equations;

iv. Variations; and

v. Change of Subject of Formulae.

The details of the domain is in Appendix A as it appears in the curriculum.

(b) Blue-Print: The Table of specification relates the content topics with the expected behaviour outcomes of the STAT. The main sub-topics have been categorized into:-

1.0 Factorization;

2.0 Equations; and

3.0 Variations.

The expected behaviour outcomes identified for STAT are:-

A.0 Knowledge;

B.0 Mathematical Skills; and

C.0 Problem-Solving.

The details of the blue-print of this study that go with this manual has been presented in Section 4.2 of Chapter 4.

(c) Item Selection and Item Statistics: Based on the blue-print, 212 items were first constructed for content validation. and first try-out. 135 items were used in this try-out to form five sets of tests. The results of the tests necessitated a workshop with the teachers on the job. The workshop was conducted at Kano Educational Resource Centre on 9th October, 1982 which discussed among other things the algebra syllabus of form 3, the text-books in use and 185 refined items from the original 212 items developed before the first try-out. A second try-out took place on 26th November, 1982 after refining the items constructed. Again 135 items were used to form five sets of tests.

The results of the second try-out were then subjected to statistical tests which helped in selecting the items for the actual tests that have been standardized in this study. The five sets of tests were found to be parallel by virtue of the equality of their means and standard deviations. It was also established that there was no significant difference between the means and standard deviations of all the five sets of tests at both 0.05 and

0.01 level of significance. This implies that statistically the tests were parallel. Table 4.3 on significance of means and variance of section 4.5 chapter 4 goes with this manual.

Other statistics done with the results of the second try-out involved the determination of the difficulty level and discriminating power of the five sets of tests. The table in Appendix E on difficulty level and discriminating power goes with this manual. These statistics also guided the selection of the actual items in the tests. Two sets of tests were prepared based on the above statistics which included 27 items each. The validity of tests was ascertained by content with the help of the teachers at the workshop and later by some other teachers of Forms 3 and 4.

(d) Normative Data:- Five different norms have been identified in this study. These are:

(i) Grade norm for total sample for each test set;

(ii) Boys' norm for each test set;

(iii) Girls' norm for each test set;

(iv) Private Schools norm for each test set; and

(v) State Government Schools norm for each test set.

All the norms are for Form 3 students in Kano State of Nigeria.

For each of the above norms the results have been reported first, by frequency distribution tables and frequency polygons for each test set; second, by percentile rank for each test set; and third by normalized standard scores for each test set. Appendix I gives the Percentile Ranks and Appendix J provides the Normalized Standard Scores. Also Figs. 5.1 - 5.10 present the Frequency Polygons for each norm.

The means and standard deviations for each norm by test set have also been presented in Table 5.1 of Chapter 5 as well as the Skewness and Kurtosis, the reliability coefficients, and standard error of measurement in Tables 5.3, 5.5 and 5.6 respectively.

All the above tables and figures will also go with this manual.

(e) Directions For Administering And Scoring The Tests:- The general instructions for taking the tests have been stated on the cover of the tests. In addition some of the test items have been grouped together and directions for answering them have been given. (See the tests in Appendix F). The tests should be

administered to form 3 students who completed learning the topics under the domain of this study.

The marking schemes for STAT is to be strictly used when scoring the students' responses. Every correct step in answering the questions has been alloted some mark(s). The marks should only be given at the steps indicated in the marking scheme. Some items have alternative responses, if a student decides to follow any of the given alternatives the students's response should be scored accordingly. The marking Schemes is in Appendix G.

The tests and the marking schemes go with this manual.

(f) Interpretation of Scores:- Any of the two sets of STAT or both can be administered to students. The students' responses should be scored according to the specifications in the marking scheme. The teacher or the appropriate authority can then decide on which of the five norms, he/she wants to categorize the students in the school. When a decision is taken then the students' raw scores can be compared with the raw scores in the Class-Interval column of the appropriate norm. If the Percentile Rank is being used, the appropriate row of the students' raw scores should be traced to the PR column that is the Percentile Rank Column, and the reading is made.

This will indicate the percentage of students each candidate is performing better than in the state.

For example if a candidate's raw score is 56% in Test A, and his school is being categorized under Private Institutions, then his Percentile Rank will be 98.9. This means that the student performs better than 98.9% of all the students in Private Institutions in Kano State, and 1.1% of the students in Private Institutions perform better than him.

On the other hand if the Normalized Standard Scores is being used then the same student with raw score of 56% in Test A is 2.3 standard deviation above the mean distribution of all the students' scores in Private Institutions in Kano State. This corresponds to the T score of 73 when traced from the Class-Interval of the raw score to the column of Z-score and T-score accordingly.

In the same manner a student under the category of Girls' Institutions with a raw score of 15 in Test B has a Percentile Rank of 86.5 and Normalized standard score of 61. This implies that the student performs better than 86.5% of all the students in Girls' Institutions in Kano State and 13.5% of the students perform better than her. Also she is 1.1 standard deviation above the mean

distribution of all the girls' scores in Girls' Institutions.

Note that a Z-score of 2.7 is the same with the T-score of 77, a Z-score of 0.6 is the same with the T-score of 56, and a Z-score of -0.4 is the same with T-score of 46. In each case both 2.7 and 77 mean 2.7 standard deviation above the mean distribution; 0.6 and 56 mean 0.6 standard deviation above the mean distribution, and -0.4 and 46 means -0.4 standard deviation below the mean distribution.

(g) The Use To Be Made of STAT:- The tests have been designed to serve two main functions. These are:

(i) Continuous Assessment. With the introduction of the New Policy on Education, the Education System demands the functional use of Continuous Assessment, especially within school system. Certain problems have been identified with the implementation process. One of them is the comparability of standards across schools. It has been suggested that one way of overcoming this problem will be the use of standardized achievement tests developed locally. The STAT can therefore be used to check the students performance across schools. It should however be noted that the

STAT is only testing achievement in one aspect of Secondary School Mathematics; and that is Algebra. For proper assessment it should be conbined with tests on other sections of secondary mathematics program which will hopefully be developed. Where school is mixed, sex norms should be used to interpret the individuals' scores.

(ii) Placement For Special Mathematics Program In Senior Secondary Schools:

The New Policy on Education also demands that there should be two stages of Senior Secondary Education. These are the Junior and Senior Secondary Schools. The Junior Secondary Schools are expected to be comprehensive and in Senior Secondary, the schools are to be categorized into specialities such as Senior Secondary School - Technical, Science, Commercial etc. For placement into Senior Secondary Schools like Science and Technical the need for sound background in Mathematics should be emphasized. The STAT can be used to ascertain the level of mathematics achievement of students before entry into such schools. The author hereby advises that the

STAT is only testing achievement in one aspect of Secondary School Mathematics; and that is Algebra. For proper assessment it should be conbined with tests on other sections of secondary mathematics program which will hopefully be developed. Where school is mixed, sex norms should be used to interpret the individuals' scores.

(ii) Placement For Special Mathematics Program In Senior Secondary Schools:

The New Policy on Education also demands that there should be two stages of Senior Secondary Education. These are the Junior and Senior Secondary Schools. The Junior Secondary Schools are expected to be comprehensive and in Senior Secondary, the schools are to be categorized into specialities such as Senior Secondary School - Technical, Science, Commercial etc. For placement into Senior Secondary Schools like Science and Technical the need for sound background in Mathematics should be emphasized. The STAT can be used to ascertain the level of mathematics achievement of students before entry into such schools. The author hereby advises that the

Percentile Rank should be used. For any candidate to be placed in Senior Secondary School Science or Technical he should have Percentile Rank of NOT less than 90 in any of the norms in this manual except the private where a Percentile Rank of 75 should be accepted.

6.3. Recommendations For Further Research.

The data obtained in this study are quite interesting. From them a number of issues for research work can be extracted. First the students' poor performances as reflected in this study can be investigated from their responses item wise. A diagnostic study can therefore be done to find out the causes of their poor performance, and suggestions made for improvement.

Second, the tests results can be subjected to further statistical analysis in order to lower the standard of achievements such that a required percentage lower than the one suggested in this study may be obtained. Item analysis such as difficulty level and discriminating power can be undertaken. Those of them proving very difficult or showing poor discriminating power can be reviewed or improved or replaced. When improvement and replacement have been made, the improved tests can be validated and the reliability should be established.

The tests can then be standardized using similar groups of this study.

Third, the test of significance of variance can be done to further investigate the consistancy of the tests. A null hypothesis such as "there is no significant relationship between Test A and Test B of STAT" can be used to further investigate the correlation between the two sets of tests.

Fourth, school by school norms can be established from the data of this study to create very local norms at either Local Government Area or Zonal Education Area Levels.

On the other hand further research work on the reliability of STAT using same group of students taking the two sets of tests should be carried out. This will allow the proper use of Pearson Product-Moment Correlation to actually determinethe extent to which the two tests are paraliel. Along with this the criterion-related validity coefficient should be established using the students scores in other tests as suggested by Brown (1983).

Assessment of the stability of STAT overtime should also be undertaken. This will help in determining the need for review of STAT for effective use.

Further research work on STAT can be undertaken to explore the possibility of using STAT in other states of Nigeria apart from Kano State, especially in Kaduna State where the first work started. This could be regarded validation exercise.

It is believed that further studies on standardization in other areas of mathematics as well as all subjects taught within our school system will provide rewarding research activities in the interest of the New National Policy on Education. It is hereby strongly recommended that all States Ministries of Education, as well as the Federal Ministry of Education and all Institutes of Education and Faculties of Education of Nigerian Universities, embark on research work on standardization of tests in all subject areas for the success of our education system.

Finally all examining bodies in the country should use parallel standardized tests, so that in any given test for any subject more than one test can be administered. This will help check examination malpractices such as leakages that keep on recurring in the country.

BIBLIOGRAPHY

Adams, G.S. & Torgerson, T.L. Measurement and Evaluation in Education Psychology and Guidance; Holt, Rinehart and Winston, 1966.

Adelaja, S.O., Lassa, P.N., Oyelese, E.A., Paling, D. & Wardle, M.E. Oxford Secondary Mathematics for Nigerian Secondary Schools 3; University Press Limited, In Association with Oxford University Press. 1982.

Ahmann, J.S. & Glock, M.D. Evaluating Pupil Growth, by Allyn and Bacon, Inc. Boston, 1963.

Airasian, P.W. & Madaus, G.J. "Functional Types of Students evaluation, Measurement and Evaluation in Guidance," 4, 1972, in N.E. Gronlund Measurement and Evaluation in Teaching; 1981.

Anastasi, A. (ed) Psychological Testing, New York, The Macmillan Company, 1959.

Ausubel, D.P. Educational Psychology A Cognitive View; Holt, Rinehart and Winston, In. 1968.

Avital, S.M. and Shettleworth, S.J. Objective for Mathematics Learning, some ideas for the Teacher. The Ontario Institute For Studies in Education, 1968.

Babcock, G.R. The Relationship Between Basal Measurement Ability and Rational Number Learning, at Three Grade Levels. Ph.D. Thesis, University of Alberta, Canada; 1978.

Begle, E.G. Critical Variables in Mathematics Education; Findings From a Survey of the Empirical Literature. Mathematical Association of America and the National Council of Teachers of Mathematics, 1979.

Bloom, B.S., Hastings, J.T. & Madaus, G.F. Handbook on Formative and Summative Evaluation of Student Learning. McGraw Hill, Book Company, New York, 1971.

Bloom, B.S. (editor) Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognative Domain; David McKay Co; 1956.

Braswell, J.S. "The College Board Scholastic Aptitude Test; An overview of the Mathematical Portion" in the Mathematics Teacher; March, 1978, Volume 71, No.3.

Brown, F.G. Principles of Educational and Psychological Testing Holt, Rinehart and Winston, 1976.

Brown, F.G. Principles of Educational and Psychological Testing 3rd Edition C.B.S. College Publishing, New York, NY 1983.

Buch, M.B. A Survey of Research in Education Baroda, The M.S. University of Baroda, 1971.

Channon, J.B., Smith, A.M., Head, H.C., Kalejaiye, A.O. & Macrae, M.F. New General Mathematics for West Africa 3; New Edition, Longman Group Ltd. 1981.

Cooney, T.J., Davis, E.J. & Henderson, K.B. Dynamics of Teaching Secondary School Mathematics. Houghton Mifflin Company, Boston; 1975.

Cronbach, L.J. Essentials of Psychological Testing. Harper and Raw, Publishers New York, Evanston, and London, 1970.

Crouse, R. & Jacobson, C. "Testing in Mathematics- What are we really measuring?" The Mathematics Teacher, November, 1975, Volume 68, No.7.

Dressel, P.L. Handbook of Academic Evaluation San Francisco; Jossey-Bass Publishers, 1976.

Ebel, R.L. Essential of Educational Measurement; Prentice-Hall Inc; Eglewood Cliffs, New Jersey, 3rd Edition, 1979.

Educational Testing Services, "ETS Builds a test" 1965 in Principles of Educational and Psychological Testing Brown, 1976.

Ema, E.B.U., Jahun, I.U., Chee, H. & Sin, K.M. New School Mathematics; Junior Secondary 2 Africana - FEP Publishers Ltd., 1985.

Ferguson, G.A. Statistical Analysis in Psychology and Education; McGraw Hill Inc. Fifth Edition, 1981.

Ferguson, R.L. & Schmeiser, C.B. "The Mathematics Usage Test of the ACT Assessment Program; An overview of its purpose, content, and use" in the Mathematics Teacher; March, 1978, Volume 71, No.3.

Further Mathematics Curriculum (Draft) Federal Ministry of Education, Lagos, 1981.

Garba, U.A. Construction and Standardization of an Achievement Test in Intergrated Science for Leavers of Junior Secondary Schools in Gongola State; Unpublished Master's Thesis. University of Maiduguri, 1984.

Garrett, H.E. Statistics in Psychology and Education; Bonbay, Vakils Feffer and Simeus Private Limited, 1960.

Glass, G.V. & Stanley, J.C. Statistical Methods in Education and Psychology; Prentice-Hall, Inc; Englewood Cliffs, New Jersey; 1970.

Glasv, R. & Nitko, A.J. "Measurement in Learning and Instruction" in R.L. Thorndike (ed) Educational Measurement 2nd ed. (Washington, D.C. American Council on Education, 1971).

Goldberg, J. & Swetz, F. "Mathematics Examinations in the Soviet Union", The Mathematics Teacher March, 1977, Volume, 70 No.3.

Good, C.V. Essentials of Educational Research Methodology and Design; Menedith Cooporation, New York, N.Y. 1972.

Good, C.V. Dictionary of Education. New York, McGraw Hill Book Company, Inc., 1945.

Goslin, D.A. Teachers and Testing. Russell. Sage Foundation, 1967.

Gronlund, N.E. Measurement and Evaluation in Teaching, Macmillan Publishing Co., Inc., New York, 1981.

Hammill, D.D. Evaluating Children for Instructional Purposes. Academic Therapy, 1971.

Hartung, M.L. "Basic Principles of Evaluation" in Evaluation in Mathematics, National Council of Teachers of Mathematics, (NCTM), 1965.

Hill, W.H. Improvement of Examinations, University of Culleut, 1972.

Holmen, M.G. & Docter, R.F. "Criticisms of Standardized Testing" in A Standardized Testing Issues of National Education Association of the United States; 1977.

Hooper, R. The Curriculum; Context, Design and Development; Oliver and Boyd Edinburgh 1977.

Hummel, T.J. "The Randomized Comparative Experiment and its Relationship to other Research Procedures". Research in Mathematics Education, National Council of Teachers of Mathematics (NCTM), 1980.

Husen, T. (ed) "International Study of Achievement in Mathematics; A Comparion of Twelve Countries". New York; Wiley 1967 2 vols. in Bloom et al Handbook on Formative and Summative Evaluation of Students Learning, 1971.

Ikeotuonye, A.I. A Validation Study of the Differential Aptitude Tests in Kaduna State of Nigeria: Implications for Guidance. Ph.D. Thesis University of Wales, 1982.

Ingenkamp, K. Developments in Educational Testing. Volumes I & II 142 University of London Press Ltd., 1969.

Jahun, I.U. "Evaluating Junior Secondary School Algebra, New Syllabus" Abacus; The Journal of the Mathematical Association of Nigeria Volume 16, 1982.

Jahun, I.U., Oyewo, J.A. & Afolabi, F.O. Handbook on Implimenting Continuous Assessment in Kano State Schools. A Kano Educational Resource Centre Publication on the Implimentation Strategy of the National Policy on Education, 1984.

Johnson, D.C. "The Research Process". Research In Mathematics Education, National Council of Teachers of Mathematics (NCTM), 1980.

Jones, C.O., Rowen, M.R. & Taylor, H.E. "An overview of the Mathematics Achievement Tests Offered in the Admissions Testing Program of the College Entrance Examination Board" The Mathematics Teacher, March, 1977, Vo. 20 No.3.

Junior Secondary School Mathematic Curriculum, Federal Ministry of Education, Lagos, 1979.

Junior Secondary School Integrated Science Core Curriculum, Federal Ministry of Education, Lagos, 1979.

Kalejaiye, A.O. "Evaluation of Achievement in Secondary School Mathematics" Abacus; The Journal of the Mathematical Association of Nigeria, Volume 16, 1982.

Kurkarni, S.S. "Constructing Equivalent Achievement Survey Tests to be used for Different Student Populations in India" Developments in Educational Testing, Volume 1, 1968

Lennon, R.T. "Norms" (1963), Testing Problems in Perspective by Ananstasi, A. 1966

Lord, F.M. "Estimating Test Reliability, Educational and Psychological Measurement"; 1955 in Ferguson, G.A. Statistical Analysis in Psychology and Education, 1981.

Marshall, J.C. & Hales, L.W. Classroom Test Construction; Addison-Wesley Publishing Company Inc., 1971.

Merwin, J.C. "Constructing Achievement Tests and Interpreting Scores," in Evaluation in Mathematics of National Council of Teachers of Mathematics (NCTM), 1965.

Mott, W. "Implementing Mathematics Proficiency Testing" The Mathematics Teacher. January, 1980, Volume 73 No.1.

Mukherjee, A. Educational Psychology; K.P. Basu Publishing Co., India, 1978.

Mukherjee, A. and Mshelia, A.Y. Test and Measurement. Manuscript for Publication, 1980.

Myers, J.L. Fundamentals of Experimental Design 3rd Edition. Allyn and Bacon, Inc; Boston, Massachusetts 02210. 1979.

Myers, S.S. "Published Evaluation Materials"; Evaluation in Mathematics, National Council of Teachers of Mathematics (NCTM), 1965.

Nair, A.S. & Soman, K. The Kerala Test of Mathematics Achievement (Cognative Domain) Standard IX Trivendrum University, Dept. of Education, 1974.

National Policy of Education, Lagos, Federal Ministry of Education. (Revised, 1981).

Ndubisi, A.F. Curriculum Objective For Effective Teaching. Africana Educational Publishers (Nig) Ltd., 1981.

Nyakutse, G.N. A Study of Influence of Home-background, Culture, Personality, Selfconcept and Attitude Factors on the Teaching Performance of Students Teachers. Ph.D. Thesis, Ahmadu Bello University, Zaria. 1985.

Ohuche, R.O. & Akeju, S.A. Testing and Evaluation in Education; African Educational Resources (AER) Printed by Skyway Press Limited, Ikeja, 1977.

Ojerinde, D. & Falayajo, W. Continuous Assessment A New Approach, University Press Limited; Ibadan, 1984.

Ojo, J.O., Ogum, G.E.O. & Sambo, A.A. MacMillam Junior Secondary Mathematics Bk. 3 MacMillan Nigeria Publishers Ltd., 1981.

Okosi, C.N., Feris, V.W. & Busbridge, J.N. Junior Secondary Mathematics. Students Book 3, Evans Brothers Ltd., 1982.

Pathan, Y. An Experimental Application of the Taxonomy of Educational Objectives in the Measurement of Cognative Processes. M.A. Dissertation University of London Institude of Education, 1969.

Popham, W.L. Creterion-Referenced Measurement. Prentice Hall, Inc. Englewood Cliffs Tersay, 07632, 1978.

Popham, W.J. Modern Educational Measurement. Prentice-Hall, Inc. Englewood Cliffs, N.J. 07632, 1981.

Quinto, F. "Teacher Made Test-An Alternative to Standardized Tests," in Standardized Testing Issues of National Education Association of United State; 1977.

Remmers, H.H., Gange, N.L. & Rumel, J.F. A Practical Introduction to Measurement and Evaluation, New Delhi Universal Book Stall, 1960.

School Mathematics Study Group. First Course in Algebra. New Haven, Conn; Yale University Press, 1961 in Bloom et al, Handbook for Formative and Summative Evaluation of Students Learning, 1971.

Sobel, M.A. & Johnson, D.A. "Analysis of Illustrative Test Items"; in Evaluation in Mathematics of National Council of Teachers of Mathematics (NCTM), 1965.

Scriven, M. "The Methodology of Evaluation", 1967; in Frederick G. Brown, Principles of Educational and Psychological Testing; Holt Rinehart and Winston, 1976.

Stanly, J.G. & Hopkins, K.D. Educational and Psychological Measurement and Evaluation; Englewood Cliffs, Prince-Hall, Inc. 1972.

Stoker, D.W. "Testing and Grading: Procedures for-Improved Students Motivation". The Mathematics Teachers, September, 1977, Volume 70, No.6.

Sueltz, B.A. "The Role of Evaluation in the Classroom" in Evaluation in Mathematics of National Council of Teachers of Mathematics (NCTM), 1965.

Suydam, M.N. Evaluation in the Mathematics Classroom. The Ohio State University Columbus, Ohio, 1974.

Swadener, M. & Wright, D.F. "Testing in the Mathematics Classroom," The Mathematics Teacher, January, 1975, Volume 68, No.1.

Thorndike, R.L. "Reliability", 1963. In Testing Problems in Perspective by Anastasi, A. 1966.

Thorndike, R.L. & Hagen, E.P. Measurement and Evaluation in Psychology and Education. 4th ed. New York; John Wiley and Sons, 1977.

Tuckman, B.W. Measuring Educational Outcomes Harcourt Brace Joravovick, Inc., 1978.

Wallace, G. & Larsen, S.C. Educational Assessment of Learning Problems; Testing for Teaching. Allyn and Bacon, Inc; 1978.

White, A.L. "Avoiding Errors in Educational Research" Research in Mathematics Education of National Council of Teachers of Mathematics (NCTM), 1980.

Wilson, J.W. "Evaluation of Learning in Secondary School Mathematics" in Bloom et al; Handbook on Formative and Summative Evaluation of Student Learning; McGraw-Hall Book Company, 1971.

Wood, R. "Objectives in the Teaching of Mathematics" Educational Research, 1968 in Bloom at el; Handbook on Formative and Summative Evaluation of Students Learning. 1971.

APPENDIX B

PART OF THE RESULTS OF THE FIRST TRY-OUT
(KADUNA STATE)

SET A

S/N.	No of Students who scored the Item Right	No of Students who scored the Item Wrong	Total
1.	-	58	58
2.	-	58	58
3.	-	58	58
4.	-	58	58
5.	-	58	58
6.	1	57	58
7.	2	56	58
8.	1	57	58
9.	2	56	58
10.	1	57	58
11.	-	58	58
12.	-	58	58
13.	-	58	58
14.	3	55	58
15.	-	58	58
16.	-	58	58
17.	4	54	58
18.	1	57	58
19.	-	58	58
20.	1	57	58
21a.	-	58	58
21b.	1	57	58
21c.	1	57	58
22.	-	58	58
23.	-	58	58
24.	-	58	58
25.	-	58	58
26a.	-	58	58
26c.	-	58	58
27.	-	58	58

SET B

S/N.	No of Students who scored the Item Right	No of Students who scored the Item wrong	Total
1.	1	52	53
2.	2	51	53
3.	-	53	53
4.	-	53	53
5.	1	52	53
6.	-	53	53
7.	3	50	53
8.	1	52	53
9.	14	39	53
10.	-	53	53
11.	1	52	53
12.	-	53	53
13.	-	53	53
14.	1	52	53
15.	1	52	53
16.	3	50	53
17.	2	51	53
18.	1	52	53
19.	1	52	53
20.	-	53	53
21a.	-	53	53
21b.	-	53	53
21c.	-	53	53
22.	2	51	53
23.	-	53	53
24.	-	53	53
25.	-	53	53
26a.	-	53	53
26b.	-	53	53
27.	-	53	53

SET C

S/N.	No of Students who scored the Item Right	No of Students who scored the Item Wrong	Total
1.	1	61	62
2.	3	59	62
3.	1	61	62
4.	2	60	62
5.	5	57	62
6.	4	58	62
7.	14	48	62
8.	3	59	62
9.	26	36	62
10.	7	55	62
11.	-	62	62
12.	-	62	62
13.	1	61	62
14.	6	56	62
15.	5	57	62
16.	-	62	62
17.	-	62	62
18.	1	61	62
19.	6	56	62
20.	-	62	62
21a.	1	61	62
21b.	-	62	62
21c.	-	62	62
22.	2	60	62
23.	-	62	62
24.	-	62	62
25.	-	62	62
26a.	-	62	62
26b.	-	62	62
27.	-	62	62

SET D

S/N.	No of Students who scored the Item Right	No of Students who scored the Item Wrong	Total
1.	-	66	66
2.	-	66	66
3.	1	65	66
4.	-	66	66
5.	7	59	66
6.	-	66	66
7.	3	63	66
8.	-	66	66
9.	-	66	66
10.	-	66	66
11.	-	66	66
12.	-	66	66
13.	-	66	66
14.	5	65	66
15.	1	65	66
16.	-	66	66
17.	-	66	66
18.	9	57	66
19.	3	63	66
20.	-	66	66
21a.	3	63	66
21b.	1	65	66
21c.	-	66	66
22.	-	66	66
23.	2	64	66
24.	-	66	66
25.	-	66	66
26a.	-	66	66
26b.	-	66	66
27.	-	66	66

APPENDIX C

STANDARDIZED TEST RESEARCH INFORMATION SHEET
(TEACHERS & SCHOOL).

Name: .....

Address: .....

Qualification: Grade II NCE/B.ED/B.SC/M.SC/Ph.D Others:....

.....

Teaching Experience: .....

Primary School:.....
Secondary School:.....

No. of Maths Teachers in your School:.....

No. of Class(es) in your School:.....

Lesson per week per Class:.....

Class(es) taught: .....

How long have you been in your present station:.....

School WASC Result (Maths)

Year	Percentage Pass
1977	.....
1978	.....
1979	.....
1980	.....
1981	.....
1982	.....

How often do you give test to your Students: .....

.....

How do you develop and administer sessional Test in your School: .....

.....

Do you use Published Tests? If yes, how effective are they?

.....

APPENDIX D

THE FIVE SETS OF TESTS FOR THE SECOND TRY-OUT.

Junior Secondary 3 Algebra Test

SET A

Name:..... Date: 26/11/82.

Instructions:

This test is aimed at assessing your total achievement in the Algebra Course of your Mathematics Curriculum. It is expected to be normalized within Kano State.

In this test each item is designed to test an important idea from this part of your Mathematics content. You do your best to answer each question correctly.

Answer each question on the space provided below it unless otherwise stated. You may do your rough work on the opposite side.

ATTEMPT ALL QUESTIONS. Time Allowed: $2\frac{1}{2}$ hours.

Set A: Page. 1

Factorize the following expressions.

1) $ck - mk$ 2) $1 - 16\,V^2$

3) Write down the expression connecting the variables in this statement:

" y varies inversely as x "

Factorize the following.

4) $q^2 - q - 6$ 5) $12\,nk + 6\,kj - 16\,jn - 8\,j^2$

6) $x^2 - 49\,k^2$ 7) $8\,xp - 2\,xp - 6qy + 36qy$

Evaluate using the idea of factors.

8) $1/4 \times 77 - 1/4 \times 31$ 9) $(1.56)^2 - (0.44)^2$

Solve the following pairs of equations by elimination.

10) $5x - 3y = 11$
$7x + 2y = 3$

11) $x + 4y = 12$
$3y - x = 16$

Set A: Page 2

Solve the following equations.

12) $\frac{n+7}{6} = \frac{n}{3}$

13) $\frac{x+3}{x^2-1} - \frac{1}{x-2} = 0$

Factorize.

14) $6xyb + 3xyq - 8kb - 4kq$

15) $5q^2 + 12q - 9$

Solve the following equations by substitution.

16) $3x + y = 21$
$x + 4y = 7$

17) $2x - y = 4$
$3x + y = 11$

Set A: Page 3

18) Make y the subject of the formula $V = \overline{II}\, r^2 - \overline{II}\, jy$.

19) Complete the following table if the first line is directly proportional to the second. Draw the graph of the table and state if y oc x. (Use graph sheet) Use 1 cm for one unit on the x-axis and 2cm for one unit on the y-axis.

x	2	3		7	
y	4/7		1 1/7		2 4/7

20) Two divided by three plus the present age of Binta is equal to one divided by her age five years ago. What is her present age?

21) The number of Naira (₦) obtainable for a number of pounds sterling (£) is given in the following table: Use 1cm to represent £1 and ₦1 on the axes respectively.

£	1	3	5	7	9
₦	1.6	4.8	8	11.2	14.4

Show graphically that ₦ oc £. Find the scale factor K. Read off the value of ₦ 10 in £ and £8 in Naira.

22) Two books and five calculators cost ₦197.00 One calculator costs ₦2 more than three books. Find the cost of one book and two calculators.

Set A: Page 4

23) If m $\propto$ n and m = 7 when n = 10 1/2, find:

(a) m when n = 7 1/2 (b) n when m = 15

24) Make h the subject of the formula $V = \overline{II}\, r^2 h$. Hence find h if V = 99 and r = 3. (Take $\overline{II} = 22/7$)

25) Solve the following simultaneous equations graphically.

$x + 2y = 5$ Take the values of x from +2 to +4
$x + y = 4$

26) $x - y = 1$ Take the values of x from 0 to +3
$2x + y = 5$

Use the graph sheets provided for Nos. 25 and 26.

27) The cost of mass production of Peugeot cars in Kaduna is partly constant and partly varies as the number of cars produced. If the total cost of making 20 cars is ₦80 000 and of making 70 cars is ₦257 600, find the cost of 200 cars.

Junior Secondary 3 Algebra Test

SET B

Name: ..... Date: 26/11/82.

Instructions:

This test is aimed at assessing your total achievement in the Algebra Course of your Mathematics Curriculum. It is expected to be normalized within Kano State.

In this test each item is designed to test an important idea from this part of your Mathematics content. You do your best to answer each question correctly.

Answer each question on the space provided below it unless otherwise stated. You may do your rough work on the opposite side.

ATTEMPT ALL QUESTIONS. Time Allowed: $2\frac{1}{2}$ hours.

Set B: Page 1

Factorize the following expressions.

1) $np + mp$ 2) $1 - j^2k^2$

3) Write down the equation connecting the variables in this statement.

" y varies directly as x "

Factorize the following.

4) $y^2 - 3y + 2$ 5) $21nz + 2zq - 7nq - 6z^2$

6) $k^2v^2 - v^2$ 7) $2mp^2 - 6m^2p$

8) Evaluate using the idea of factors.

$(108)^2$ 9) $(4.4)^2 - (2.6)^2$

Solve the following pairs of equations by elimination.

10) $y - 2x = 4$
$y + x = 7$

11) $5x + 3y = 1$
$2x + 3y = -5$

Set B: Page 2.

Solve the following equations.

12) $\frac{x}{2} - \frac{x+1}{4} = 3$ 13) $\frac{2}{a+2} = \frac{3}{a-3}$

Factorize.

14) $6up - 2upk - 9rv + 3krv$ 15) $4y^2 + 8y + 3$

Solve the following equations by substitution.

16) $x + y = 2$
$2y - x = -13$

17) $y - 2x = 1$
$2x - 4y = 1$

18) Make S the subject of the formula $M = \sqrt{\frac{0 - 1/4\ S}{k}}$

Set B: Page. 3

19) Complete the following table if the first line is directly proportional to the second. Draw the graph of the table and state if y ∝ x. (Use graph sheet) Use 2cm for one unit on the x-axis and 1cm for $\frac{1}{3}$ unit on the y-axis.

x	5	6	7		9
y		2		2 2/3	

20) When seven is subtracted from a number and the difference is divided by three, the result is the same as adding five to the number and dividing the sum by two. What is the number?

21) Draw the graph of the table below.

(a)

x	1	2	3	4	5	6
y	5	10	15	20	25	30

(b) Read off the value of:

(i) x when y = 17 ———————

(ii) y when x = 3 ———————

(iii) x when y = 29 ———————

(iv) y when x = 5.3 ———————

Note: Use 2cm for one unit on the x-axis and 1cm for 5 units on y-axis.

22) The sum of two numbers is 19 and their difference is 5. Find the numbers.

Set B: Page 4.

23) If $V \propto \frac{1}{r^2}$ and $V = 13$ when $r = 3$; find the scale factor k. Find also V when $r = 5$ and r when $V = 9$.

24) Find n if $B = 132$, $r = 3$ from $B = 2\overline{II}rn + 2\overline{II}r^2$

. (Take $\overline{II} = 22/7$)

Solve the following simultaneous equations graphically using the graph sheets provided.

25) $x + y = 3$
$2x - 5y = -1$
Take values of x from +1 to +3

26) $2x + y = 2$
$x - 2y = 1$
Take values of x from 0 to +3

27) The time required for a bus ride into the country varies directly as the distance and inversely as the speed. If a trip of 84 km takes 2 hours and 24 minutes, how long would it take (under similar conditions) to make a trip of 189 km?

Junior Secondary 3 Algebra Test

SET C

Name: ..... Date: 26/11/82

Instructions:

This test is aimed at assessing your total achievment in the Algebra Course of your Mathematics Curriculum. It is expected to be normalized within-Kano State.

In this test each item is designed to test an important idea from this part of your Mathematics content. You do your best to answer each question correctly.

Answer ~~each~~ question on the space provided below it ~~unless~~ otherwise stated. You may do your rough work on the opposite side.

ATTEMPT ALL QUESTIONS. Time Allowed: $2\frac{1}{2}$ hours.

Set C: Page 1.

Factorize the following expressions.

1) $5v^2u - 3vu^2$ 2) $n^2-x^2y^2$

3) Write down an equation of u $\propto \frac{1}{m}$.

Factorize the following.

4) $x^2-7x + 10$ 5) $6nb^2-12b - 3nb + 6$

6) $9n^2-1$ 7) $k^2+k^2x^2-4x^2k^2+4k^2$

Evaluate using the idea of factors.

8) $3/5 \times 6/11 + 6/11 \times 2/5$ 9) $(143)^2-(57)^2$

Set C: Page 2.

Solve the following pairs of equations by elimination.

10) $x + 5y = 4$
$11x + 6y = 5$

11) $x + 2y = 5$
$x + y = 4$

Solve the following equations.

12) $\frac{2x-3}{4} + \frac{6-2x}{5} = 1$

13) $\frac{3}{4x+3} = \frac{1}{x-7}$

Factorize.

14) $6xz + 3qz - 10xy - 5yq$

15) $2x^2 - 3x - 5$

Solve the following equations by substitution.

16) $x + y = 4$
$2x - y = 5$

17) $5x - y = 6$
$2x + 3y = 16$

Set C: Page 3.

18) Make θ the subject of the formula $U = \frac{\theta}{270} \overline{II}\ h'$

19) Complete the following table if the first line is directly proportional to the second. Draw the graph of the table and state if y oc x. (Use graph sheet) Use 1cm for one unit on the x-axis and 1cm for 5 units on the y-axis.

x	2	3	5		9
y		15		35	

20) When four is added to one-fifth of a number, the result is the same as adding six to two-thirds of the number. What is the number?

21) x varies as given in the equation $4x = y-6$. Use this equation to complete the table below.

x	0	1	2	3	4	5	6
y							

Plot the graph of x against y. Read off the value of:

(i) x when y = 20 - - - - -

(ii) y when x = 4 1/2 - - - - -

Note: Use 1 cm for one unit on the x-axis and 1cm for 5 units on the y-axis.

22) Hamdala Hotel is 6 metres higher than Durba Hotel. Find the height of Hamdala Hotel if the sum of their heights is 28 m.

Set C: Page 4.

23) If $x \propto m$ and $x = 15$ when $m = 6$, find:

(a) x when m = 4 (b) m when x = 65

24) Find r if U = 99 from $U = 7/2\ \pi r^2$ (Take $\pi = 22/7$)

Solve the following simultaneous equations graphically. Use the graph sheets provided.

25) $x - 3y = 6$
$x + 3y = 0$
Take the values of x from +2 to +4

26) $2x + y = 5$
$x - y = 4$
Take the values of x from +1 to +4

27) The distance D covered in a certain journey is directly proportional to the time T taken to do the journey and directly proportional to the age A of the vehicle used for the journey.

(a) Write an equation for this variation

(b) When the time T is 4 hours and the scale factor k is 7, the distance D is 140 km. Find the age A of the car.

Junior Secondary 3 Algebra Test

SET D

Name: ..... Date: 26/11/82

Instructions.

This test is aimed at assessing your total achievement in the Algebra Course of your Mathematics Curriculum. It is expected to be normalized within Kano State.

In this test each item is designed to test an important idea from this part of your Mathematics content. You do your best to answer each question correctly.

Answer each question on the space provided below it unless otherwise stated. You may do your rough work on the opposite side.

ATTEMPT ALL QUESTIONS. Time Allowed: 2½ hours.

Set D: Page 1.

Factorize the following expressions.

1) $yq^2 - y^2q$

2) $1 - v^4$

3) Write down $V \propto \frac{1}{n}$ in words.

Factorize the following:

4) $n^2 - 2n - 3$

5) $20kv - 12vl - 35kj + 21lj$

6) $49n^2 - b^2$

7) $6n^3 - 4n^2 + 2n$

Evaluate using the idea of factors.

8) $36 \times 19 - 36 \times 9$

9) $(95)^2$

Solve the following pairs of equations by elimination.

10) $x - 2y = -2$
$2x - 3y = -1$

11) $6x - 2y = 9$
$3x + 4y = 12$

Set D: Page 2.

Solve the following equations.

12) $\frac{x+1}{4} + \frac{x+5}{2} = 5$ 13) $\frac{3}{2n+4} - \frac{2}{n+2} = \frac{2}{n+2}$

Factorize.

14) $kv - ku - mv + mu$ 15) $3n^2 + 2n - 21$

Solve the following equations by substitution.

16) $y - x = 1$
$x + y = 3$

17) $2x + 2y = 2$
$x - 3y = -11$

18) Make $\overline{II}$ the subject of the formula.

$kv^2 = j\ (3n + \overline{II})$

Set D: Page 3.

19) Complete the following table if the first line is directly proportional to the second. Draw the graph of the table and state if y ∝ x. (Use graph sheet and use 1cm to represent 2 units on the x-axis and 1cm to represent 10 units on the y-axis).

x	3	5	8		15
y			56	77	

20 If the sum of a number and seven is divided by four, the result is the same as subtracting five from the number and dividing the difference by three. What is the number?

21) The number of Dollars ($) obtainable for a number of Naira (₦) is given in the following table

₦	2	4	6	8	10
$	3	6	9	12	15

Show graphically that $ ∝ ₦. Find the scale factor k. Read off the value of $5 in ₦ and ₦9 in $. (Use 1cm to represent One Dollar ($) and One Naira (₦) respectively).

22) Mustapha has ₦40.00 more than what Ladi has and the sum of their money is ₦150.00 How much does Ladi have?

Set D: Page 4.

23) U $\propto$ YZ. When Z = 8 and Y = 4, U = 64. Find the scale factor k. Find also U when Z = 5 and Y = 10.

24) Fine Q if m = 11, U = 55, T = 11 from $\frac{m}{1/2Q} = \frac{T\ \overline{II}}{U}$

Solve the following simultaneous equations graphically

25) $3x + 2y = 9$
$2x + y = 4$
(Take the values of x from -2 to 0)

26) $3x + y = 21$
$x + 4y = 7$
(Take the values of x from +6 to +8).

(Use the graph sheets provided).

27) 15 men can dig a ditch in 4 days. How long will it take 6 men to dig a similar ditch if they work at the same rate?

Junior Secondary 3 Algebra Test

SET E

Name:..... Date: 26/11/82

Instructions:

This test is aimed at assessing your total achievement in the Algebra Course of your Mathematics Curriculum. It is expected to be normalized within Kano State.

In this test each item is designed to test an important idea from this part of your Mathematics content. You do your best to answer each question correctly.

Answer each question on the space provided below it unless otherwise stated. You may do your rough work on the opposite side.

ATTEMPT ALL QUESTIONS. Time Allowed: $2\frac{1}{2}$ hours.

Set E: Page 1.

Factorize the following expressions.

1) $5bj + 10kj$

2) $9j^2 - 1$

3) Write down y oc x in words.

Factorize the following.

4) $m^2 - 7m + 12$

5) $18nb^2 - 12b - 18nb + 12$

6) $36x^2 - y^2$

7) $2q^4p^2 + 8q^2p^4$

Evaluate using the idea of factors.

8) $13 \times 3 + 13 \times 7$

9) $(73)^2 - (27)^2$

Solve the following pairs of equations by elimination.

10) $4x - 3y = 1$
$x - 2y = 4$

11) $5x - y = 6$
$2x + 3y = 16$

Set E: Page 2.

Solve the following equations.

12) $\frac{k+3}{5} = \frac{k+16}{7}$

13) $\frac{12}{x-4} = \frac{8}{x-5}$

Factorize.

14) $ax - ay - bx + by + cy - cx$

15) $3m^2 - 5m - 2$

Solve the following equations by substitution.

16) $x + 2y = 5$
$x + y = 4$

17) $3x - 2y = -2$
$x - 3y = 4$

18) Make d the subject of the formula $N = \frac{1}{4} d (a + b)$.

Set E: Page 3.

19) Complete the following table if the first line is directly proportional to the second. Draw the graph of the table and state if y ∝ x. (Use graph sheet and use 1cm to represent one on the x-axis and 1cm to represent $\frac{1}{3}$ on the y-axis)

x	4	5	6	7	8
y			$1\frac{2}{3}$		

20) Five divided by the difference of a number and four in that order, is the same as dividing eight by the number and five. What is the number?

21) Complete the following table for a train travelling for 120 km.

Speed (km/h	40	50	60	70	80	90	100	110	120
Time	3	2.4	2	1.7	1.5	1.3	1.2	1.1	1
$\frac{1}{\text{Time}}$									

Draw the graph of speed against $\frac{1}{\text{Time}}$. Is speed ∝ $\frac{1}{\text{Time}}$?

(Use graph sheet)

22) Three times Zainab's age plus thrice her brother's age totals 44 years. In 6 years time their ages will add up to 30 years. How old are they?

Set E: Page 4.

23) $x \propto \sqrt{y}$ and $x = 5\frac{1}{2}$ when $y = 9$; find the scale factor k. Find also x when $y = 16$.

24) If 462 cm^3 is the volume of a cone whose radius is 3 cm, what is the height of the cone? (Volume of a cone $= \frac{1}{3} \pi r^2 h$; Take $\pi = 22/7$).

Solve the following simultaneous equations graphically. Use the graph sheets provided.

25) $x + y = 5$
$2x - y = 1$
Take values of x from +1 to +3

26) $3x + 2y = 10$
$4x - y = 6$
Take values of x from +1 to +3.

27) A bag of rice feeds a group of 12 students for 7 days. How long can the bag feed 14 students?

APPENDIX E

ITEM ANALYSIS.

Difficulty Level And Discrimination Power.

S/NO	P_1	D_1	P_2	D_2	P_3	D_3	P_4	D_4	P_5	D_5
1.	.62*	.76*	.72*	.82*	.41	.70	.38	.63	.46	.72
2.	.46*	.80*	.40	.70	.37	.73	.44*	.61*	.36	.83
3.	.27*	.52*	.20	.48	.20	.68	.19	.37	.30*	.69*
4.	.33	.80	.38	.90	.48*	.73*	.38	.71	.42*	.58*
5.	.43*	.96*	.30*	.65*	.13	.53	.25	.61	.20	.58
6.	.51*	.84*	.12	.10	.44	.78	.42	.68	.45*	.86*
7.	.07	.20	.48*	.73*	.11	.10	.26	.53	.31*	.72*
8.	.37	.84	.31	.58	.40*	.65*	.39	.61	.52*	.81*
9.	.38*	.60*	.30	.53	.25	.80	.27	.45	.41*	.56*
10.	.32*	.80*	.28*	.73*	.14	.25	.24	.61	.22	.58
11.	.28*	.84*	.27*	.73*	.25	.60	.24	.58	.31	.72
12.	.51*	.96*	.15	.43	.25	.75	.23	.55	.37*	.89*
13.	.24	.56	.20	.50	.25*	.68*	.09	.16	.27*	.72*
14.	.48*	.96*	.40*	.73*	.36	.68	.39	.74	.28	.58
15.	.28	.80	.34*	.68*	.31*	.70*	.28	.63	.22	.61
16.	.27*	.56*	.10	.18	.21	.50	.24	.53	.27*	.61*
17.	.25*	.48*	.10	.18	.15	.45	.11	.32	.21*	.50*
18.	.18*	.40*	.10	.28	.22*	.68	.15	.39	.14	.42
19.	.09	.20	.17	.33	.34*	.50	.18	.05	.26*	.42
20.	.05	.08	.16*	.45*	.11	.33	.16*	.42*	.11	.36
21.	.13*	.44*	.29*	.50*	.12	.23	.07	.03	.05	.03
22.	.12	.32	.32*	.78*	.23	.60	.26*	.42*	.05	.08
23.	.22*	.72*	.09	.30	.18	.53	.22*	.53*	.13	.36
24.	.31*	.72*	.09	.28	.13*	.45*	.09	.26	.12	.31
25.	.14*	.44*	.10	.28	.07	.20	.08	.26	.11*	.31*
26.	.16*	.48*	.09*	.25*	.05	.18	.07	.24	.09	.19
27.	.02	.00	.05	.08	.12	.28	.31*	.74*	.36	.63
N	94		147		151		139		128	

*The corresponding items selected.
Pn = Difficulty level Dn = Discriminating Power.

APPENDIX F

FINAL TESTS.

Junior Secondary 3 Algebra Test

SET A

Name: ..... Date: _____

Instructions:

This test is aimed at assessing your total achievement in the Algebra Course of your Mathematics Curriculum. It is expected to be normalized within Kano State.

In this test each item is designed to test an important idea from this part of your Mathematics content. You do your best to answer each question correctly.

Answer each question on the space provided below it unless otherwise stated. You may do your rough work on the opposite side.

ATTEMPT ALL QUESTIONS. Time Allowed: 2½ hours.

Set A: Page 1.

Factorize the following expression.

1) $np + mp$

2) $1 - x^2$

3) Write down the expression connecting the variables in this statement.

"y varies inversely as x"

Factorize the following.

4) $M^2 - 7M + 12$

5) $21nz + 2zq - 7nq - 6z^2$

6) $36x^2 - y^2$

7) $2mp^2 - 6m^2p$

Evaluate using the idea of factors.

8) $13 \times 3 + 13 \times 7$

9) $(1.56)^2 - (0.44)^2$

Solve the following pairs of Equations by elimination.

10) $y - 2x = 4$
$y + x = 7$

11) $x + 4y = 12$
$3y - x = 16$

Set A: Page 2.

Solve the following equations.

12) $\frac{n+7}{6} = \frac{n}{3}$

13) $\frac{3}{4x+3} = \frac{1}{x-7}$

Factorize.

14) $6np - 2upk - 9rv + 3krv$

15) $4y^2 + 8y + 3$

Solve the following equations by substitution.

16) $x + 2y = 5$
$x + y = 4$

17) $3x - 2y = -2$
$x - 3y = 4$

18) Make θ the subject of the formula.

$U = \frac{\theta}{270} \overline{II} h.$

Set A: Page 3

19) Complete the following table if the first line is directly proportional to the second. Draw the graph of the table and state if y ∝ x. (Use graph sheet. Use 1cm for one unit on the x-axis and 1cm for 5 units on the y-axis.

x	2	3	5		9
y		15		35	

20) If the sum of a number and seven is divided by four, the result is the same as subtracting five from the number and dividing the difference by three. What is the number?

21) Draw the graph of the table below.

(a)

x	1	2	3	4	5	6
y	5	10	15	20	25	30

Use 2cm for one unit on the x-axis and 1cm for 5 units on y-axis.

(b) Read off the value of:

i. x when y = 17 .....

ii. y when x = 3 .....

iii. x when y = 29 .....

iv. y when x = 5.3 .....

22) Mustapha has ₦40.00 more than what Hafsatu has and the sum of their money is ₦150.00. How much does Hafsatu have?

Set A: Page 4.

23) If m oc n and m = 7 when n = $10\frac{1}{2}$, find:-

(a) m when n = $7\frac{1}{2}$ (b) n when m = 15

24) Find r if U = 99 from U = $\frac{7}{2}\ \overline{\text{II}}r^2$

(Take $\overline{\text{II}} = \frac{22}{7}$)

Solve the following simultaneous equations graphically. Use the graph sheets.

25) $x + y = 5$
$2x - y = 1$ (Take values of x from + 1 to +3)

26) $x - y = 1$
$2x + y = 5$ (Take the values of x from 0 to +3)

27) 15 men can dig a ditch in 4 days. How long will it take 6 men to dig a similar ditch if they work at the same rate?

Junior Secondary 3 Algebra Test

SET B

Name:..... Date: _____

Instruction:

This test is aimed at assessing your total achievement in the Algebra Course of your Mathematics Curriculum. It is expected to be normalized within Kano State.

In this test each item is designed to test an important idea from this part of your Mathematics content. You do your best to answer each question correctly.

Answer each question on the space provided below it unless otherwise stated. You may do your rough work on the opposite side.

ATTEMPT ALL QUESTIONS. Time Allowed: $2\frac{1}{2}$ hours.

Set B: Page 1.

Factorize the following expressions.

1) $CK - MK$

2) $1 - 16v^2$

3) Write down y oc x in words.

Factorize the following.

4) $x^2 - 7x + 10$

5) $12nk + 6kj - 16jn - 8j^2$

6) $x^2 - 49k^2$

7) $2q^4p^2 + 8q^2p^4$

Evaluate using the idea of factors.

8) $\frac{3}{5} \times \frac{6}{11} + \frac{6}{11} \times \frac{2}{5}$

9) $(73)^2 - (27)^2$

Solve the following pairs of equations by elimination.

10) $5x - 3y = 11$
$7x + 2y = 3$

11) $5x + 3y = 1$
$2x + 3y = -5$

Set B: Page 2.

Solve the following equations.

12) $\frac{k + 3}{5} = \frac{k + 16}{7}$

13) $\frac{12}{x - 4} = \frac{8}{x - 5}$

Factorize.

14) $6xyb + 3xyq - 8kb - 4kq$

15) $2x^2 - 3x - 5$

Solve the following equations by substitution.

16) $3x + y = 21$

$x + 4y = 7$

17) $2x - y = 4$

$3x + y = 11$

18) Make y the subject of the formula.

$v = \overline{II}r^2 - \overline{II}\,jy$

Set B: Page 3.

19) Complete the following table if the first line is directly proportional to second. Draw the graph of the table and state if y oc x. (Use graph sheet - 1 cm to represent one on the x-axis and 1cm to represnt $\frac{1}{5}$ on the y-axis).

x	4	5	6	7	8
y		$1\frac{2}{5}$			

20) When seven is subtracted from a number and the difference is divided by three, the result is the same as adding five to the number and dividing the sum by two. What is the number?

21) The number of Naira (₦) obtainable for a number of pounds sterling (£) is given in the following table:

£	1	3	5	7	9
₦	1.6	4.8	8	11.2	14.4

Show graphically that ₦ oc £. Find the Scale factor K. Read off the value of ₦10 in £ and £8 in Naira. (Use 1cm to represent £1 and ₦1 on the axes respectively).

22) The sum of two numbers is 19 and their difference is 5. Find the number.

Set B: Page 4.

23) U $\propto$ YZ when Z = 8 and Y = 4, U = 64. Find the Scale factor K. Find also U when Z = 5 and Y = 10.

24) Make h the subject of the formula from $V = \overline{II}r^2h$.

Hence find h if V = 99 and r = 3 (Take $\overline{II}$ $\frac{22}{7}$).

Solve the following simultaneous equations graphically. Use graph sheets.

25) $x + 2y = 5$
$x + y = 4$

Take the values of x from +2 to +4

26) $2x + y = 2$
$x - 2y = 1$

Take the values of x from 0 to +3.

27) A bag of rice feeds a group of 12 students for 7 days. How long can the bag feed 14 students?

APPENDIX G

MARKING SCHEME FOR THE FINAL TESTS

SET A.

1. $p(n+m)$ ..... 2 Marks.

2. $(1+x)(1-x)$ .....2 "

3. $y \propto \frac{1}{x}$ ..... 2 "

4. $(M-3)(M-4)$ ..... 3 Marks

5. $7n(3z-q) + 2z(q-3z)$ ..... (1)
   $= 7n(3z-q)-2z(3z-q)$ ..... (1)
   $= (7n-2z)(3z-q)$ ..... (1) 3 Marks.

6. $(6x+y)(6x-y)$ .....2 Marks

7. $2mp(p-3m)$ ..... 2 "

8. $13(3+7)$ ..... (1)
   $= 13 \times 10$ ..... (1)
   $= 130$ ..... (1) 3 Marks.

9. $(1.56+0.44)(1.56-0.44)$ ..... (1)
   $= 2(1.12)$ ..... (1)
   $= 2.24$ ..... (1) 3 Marks.

10. $y - 2x = 4$ ..... (1)
    $y + x = 7$ ..... (2)
    (1)-(2) .....(1)
    $-3x = 4-7$ .....(1)
    $= -3$
    $x = -\frac{-3}{3}$
    $x = 1$ ..... (1)
    Subst. in (2)
    $y + 1 = 7$
    $y = 7 - 1$
    $= 6$ ..... (1) 4 Marks.
    (1,6)

Set A: 2

11. $x + 4y = 12$ .....(1)..

$3y - x = 16$ .....(2)..

(1)+(2) ..... (1)

$7y = 28$ ..... (1)

$y = 4$ ..... (1)

Subst. in (2)

$3.4 - x = 16$

$-x = 16-12$

$-x = 4$

$x = -4$ ..... (1) 4 Marks.

$(-4,4)$ .....

12. $3n + 21 = 6n$ ..... (1)

$6n - 3n = 21$ ..... (1)

$3n = 21$ ..... (1)

$n = \frac{21}{3}$

$= 7$ ..... (1) 4 Marks

OR $\frac{n+7}{6} - \frac{n}{3} = 0$ .....((1)

$\frac{n+7}{6} - \frac{2n}{6}$

$\frac{n+7-2n}{6} = 0$ ..... (1)

$n + 7 - 2n = 0$ ..... (1)

$-n + 7 = 0$

$-n = -7$

$n = 7$ ..... (1) 4 Marks.

Set A: 3.

13. $3x - 21 = 4x + 3$ ..... (2)
$4x - 3x = -21 - 3$ ..... (1)
$x = -24$ ..... (1) 4 Marks.

OR

$\frac{3}{4x+3} - \frac{1}{x-7} = 0$ ..... (1)

$\frac{3}{4x+3} - \frac{1}{x-7}$

$\frac{3x-21 - 4x-3}{(4x+3)(x-7)}$

$\frac{3x-21 - 4x-3}{(4x+3)(x-7)} = 0$ ..... (1)

$3x-21 - 4x-3 = 0$ ..... (1)

$-x - 24 = 0$

$-x = 24$

$x = -24$ ..... (1) 4Marks

14. $2up(3-k)-3rv(3-k)$ ..... (2)
$=(2up-3rv)(3-k)$ ..... (1) 3 Marks

15. $(2y+3)(2y+1)$ ..... 3 Marks.

16. $x+2y = 5$ .....(1)
$x+y = 4$ .....(2)

From (2)

$x = 4-y$ ..... (2)

Subst. in (1)

$4-y+2y = 5$
$4+y = 5$
$y = 5-4$
$y = 1$ .....(1)

Subst. y in (2)

$x+1 = 4$
$x=4-1$
$=4$ ..... (1) 4 Marks

(3,1)

Set A: 4.

17. $3x - 2y = -2$ ..... (1)
$x - 3y = 4$ ..... (2)

From (2)

$x = 4+3y$ ..... (2)

Subst. in (1)

$3(4+3y)-2y = 2$
$12+9y - 2y = 2$
$12 + 7y = -2$
$7y = -2-12$
$= -14$
$y = \frac{-14}{7}$
$= -2$ ..... (1)

Subtst. y in (1)

$3x-2.(-2) = -2$
$3x+4 = -2$
$3x = -2-4$
$= -6$
$x = \frac{-6}{3}$
$= -2$ ..... (1) 4 Marks.
$(-2,-2)$

18. $270u = 0\overline{11}h$. ..... (1)

$\frac{0\overline{11}h}{11h} = \frac{270u}{11h}$

$\therefore \theta = \frac{270u}{11h}$ ..... (1) 2 Marks.

20. Let the number be x

then $\frac{x+7}{4} = \frac{x-5}{3}$ ..... (2)

$\therefore$ $3x+21 = 4x-20$ ..... (1)
$4x-3x = 21+20$ ..... (1)
$x = 41$ ..... (1) 5 Marks.

OR

Set A: 5.

Let the number be x

then $\frac{x+7}{4} = \frac{x-5}{3}$ ..... (1)

$\frac{x+7}{4} - \frac{x-5}{3} = 0$

$\frac{x+7}{4} - \frac{x-5}{3}$

$\frac{3x+21 - 4x+20}{12}$ ..... (1)

$\frac{3x+21 - 4x+20}{12} = 0$ ..... (1)

$3x+21-4x+20 = 0$ ..... (1)

$-x+41 = 0$

$-x = -41$

$x = 41$ ..... (1) 5Marks.

22. Let Mustapha's money be M and Hafsatu's be H

Then $M = H-40$

i.e. $M-H = 40$ .....(1) ..... (1)

and $M+G = 150$.....(2) ..... (1)

(1)+(2) $2M = 190$

$M = 95$ ..... (1)

Subst. M in (1)

$95-H = 40$

$H = 95 - 40$

$= 55$ ..... (1)

∴ Hafsatu has ₦55.00 ..... (1) 5 Marks.

OR

By substitution 5 Marks

OR

Graphically 5 Marks.

Set A: 6.

23. $M = km$

$7 = \frac{21k}{2}$

$K = \frac{1}{\cancel{7}\times\cancel{2}} \atop \cancel{21}_{3}$

$= 2/3$ ..... (1)

(a) $M = 2/3 \times 15/2$

$= 5$ ..... (1)

(b) $15 = 2/3n.$

$n = 3/2 \times 15$

$= \frac{45}{2}$

$n = 22\frac{1}{2}$ ..... (1) 3 Marks.

24. $u = 7/2\Pi r^2$

$r^2 = 2/7\frac{u}{\Pi}$ ..... (1)

$= \frac{2}{7} \times 99 \times \frac{7}{22}$ ..... (1)

$r^2 = 9$

$r = \sqrt{9}$

$= 3$ ..... (1) 3 Marks.

Set A: 7

27. As the number of men decreases the time taken to dig the ditch increases

Let M = men

T = Time

K = constent

$M \propto \frac{1}{T}$ ..... (2)
$M = \frac{k}{T}$ ..... (1) } 3

i.e. $5 = \frac{k}{4}$ ..... (1)

$k = 20$ ..... (1)

$\therefore \quad 6 = \frac{60}{T}$ ..... (1)

$T = \frac{60}{6}$

$= 10$ ..... (1)

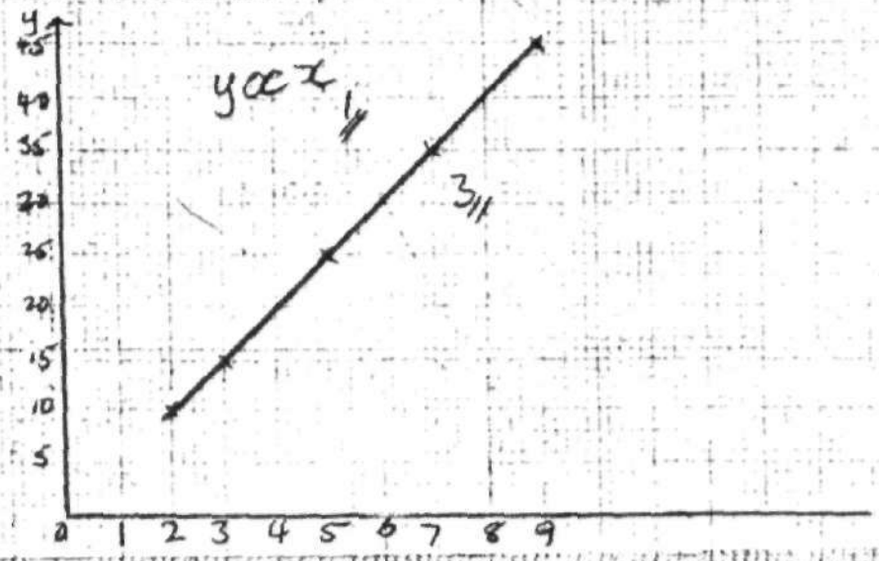
$\therefore$ Time taken by 6 men is 10 days (1) 8 Marks.

SET A

A.

x	2	3	5	7	9
y	10	15	25	35	45

2//



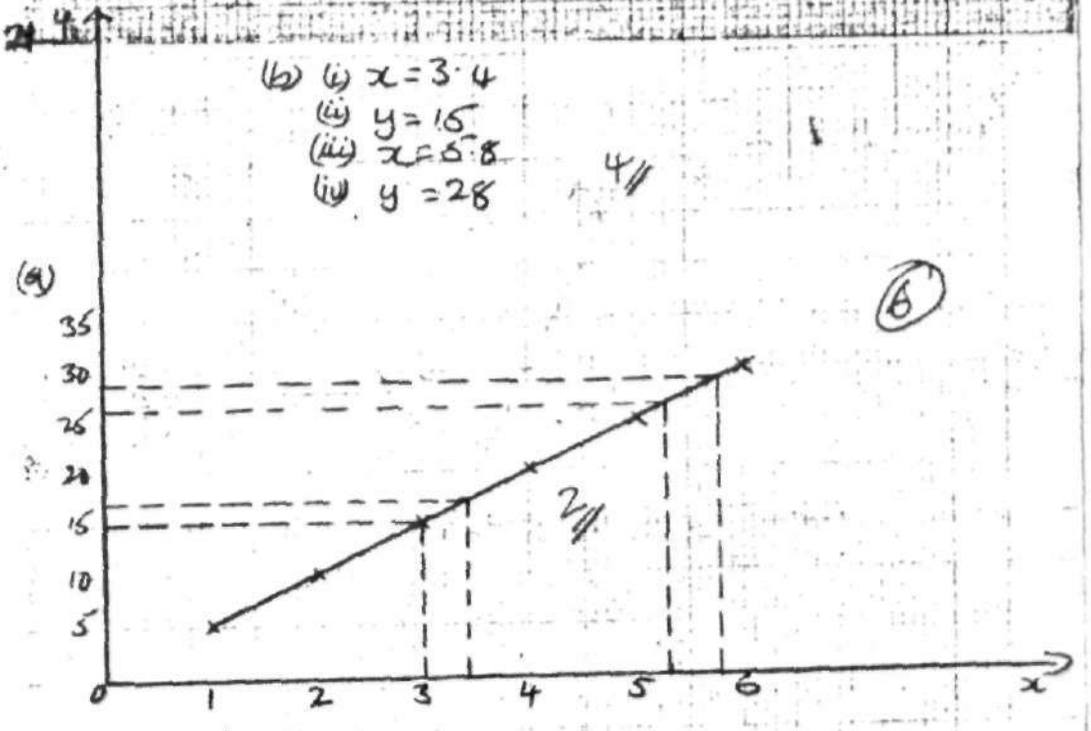
$y \propto x$ 1//

3//

(b) (i) $x = 3 \cdot 4$
(ii) $y = 15$
(iii) $x = 5 \cdot 8$
(iv) $y = 28$

4//

(a)



2//

6

SET A

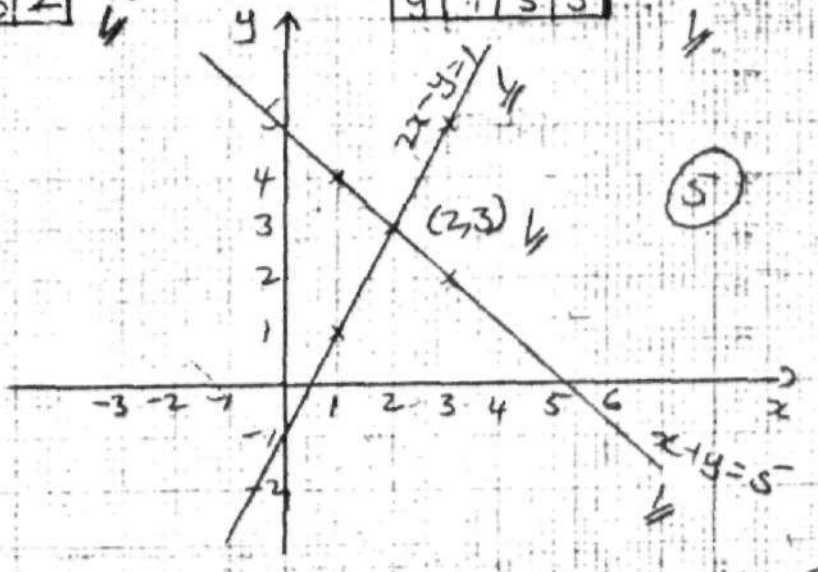
25

x	1	2	3
y	4	3	2

$x+y=5$

x	1	2	3
y	1	3	5

$2x-y=1$



5

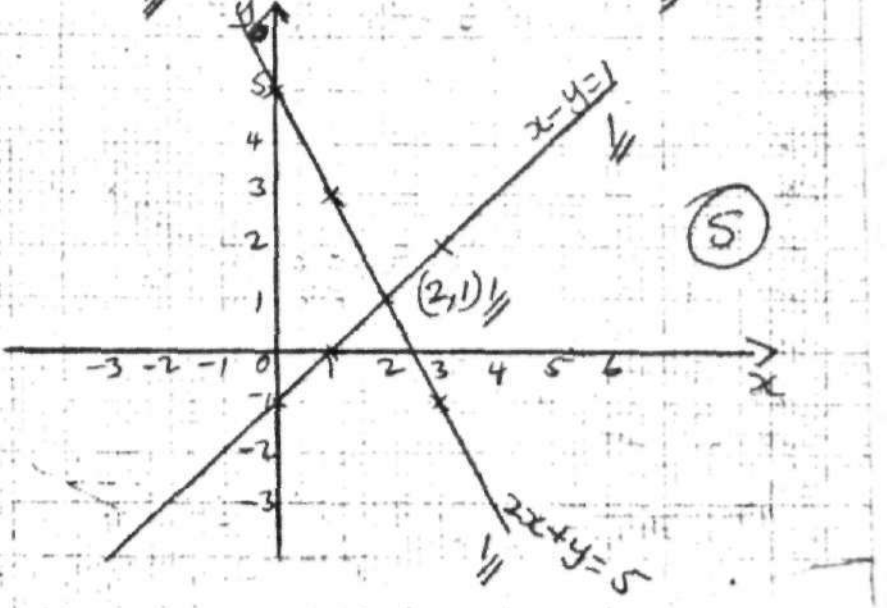
26

x	0	1	2	3
y	-1	0	1	2

$x-y=1$

x	0	1	2	3
y	5	3	1	-1

$2x+y=5$



5

SET B. 1.

1. $K(C-M)$ ..... 2 Marks.

2. $(1-4v)(1+4v)$ ..... 2 Marks.

3. y is directly proportional to x ... 2 Marks

4. $(x-2)(x-5)$ ..... 3 Marks.

5. $6k(2n+j) - 8j(2n+j)$ ..... (1)
$=(6k-8j)(2n+j)$ ..... (1)
$= 2(3k-4j)(2n+j)$ ..... (1) 3 Marks.

OR
$2[3k(2n+j) - 4j(2n+j)]$ ..... (1)
$= 2[(3k-4j)(2n+j)]$ ..... (1)
$= 2(3k-4j)(2n+j)$ ..... (1) 3 Marks.

6. $(x-7k)(x+7k)$ ..... 2 Marks.

7. $2p^2q^2(q^2+p^2)$ ..... 2 Marks.

8. $6/11\ (3/5 + 2/5) = 6/11 \times 5/5$ ..... (1)
$=6/11 \times 1$ ..... (1)
$= 6/11$ ..... (1) 3 Marks.

9. $(73 + 27)(73 - 27)$ ..... (1)
$= 100 \times 46$ ..... (1)
$= 4600$ ..... (1) 3 Marks.

10. $5x-3y = 11$ ..... (1)
$7x + 2y = 3$ ..... (2)

(1) x 2 and (2) x 3 ..... (1)
$10x-6y = 22$ .....(3) } ....(1)
$21x+6y = 9$ .....(4) }

(3)+(4) $31x = 31$
$x = \frac{31}{31} = 1$ .....(1)

Substitute x in (1)
$5x1-3y = 11$
$-3y = 11-5$
$-3y = 6$
$y = -6/3$
$=-2$ ..... (1) 4 Marks.

(1,-2).

Set B: 2.

11. $5x+3y = 1$ .....(1)
$2x+3y = -5$ .....(2)

(1)-(2) ..... (1)
$3x = 6$ ..... (1)
$x = 6/3$
$= 2$ ..... (1)

Substitute x in (2)

$2x2 + 3y = -5$
$4+3y = -5$
$3y = -9$
$y = -3$ ..... (1)
$(2,-3)$ 4 Marks.

12. $7k+21 = 5k+80$ ..... (1)
$7k-5k = 80-21$ ..... (1)
$2k = 59$ ..... (1)
$k = 29\frac{1}{2}$ ..... (1) 4 Marks.

OR

$\frac{k+3}{5} - \frac{K+16}{7} = 0$ ..... (1)

$\frac{k+3}{5} - \frac{k+16}{7} = 0$ ..... (1)

$\frac{7k+21-5k-80}{35}$

$\frac{7k+21-5k-80}{35} = 0$

$7k+21-5k-80 = 0$ ..... (1)
$2k-59 = 0$
$2k = 59$
$k = 29\frac{1}{2}$ ..... (1) 4 Marks.

13. $12x-60 = 8x-32$ ..... (1)
$12x-8x = -32+60$ ..... (1)
$4x = 28$ ..... (1)
$x = \frac{24}{4}$
$= 7$ ..... (1) 4 Marks.

OR

Set B: 3.

$\frac{12}{x-4} - \frac{8}{x-5} = 0$ ..... (1)

$\frac{12}{x-4} - \frac{8}{x-5}$

$\frac{12x-60 - 8x+32}{(x-4)(x-5)}$

$\frac{12x-60 - 8x+32}{(x-4)(x-5)} = 0$ ..... (1)

$12x-60-8x+32 = 0$ ..... (1)

$4x - 28 = 0$

$x = \frac{28}{4}$

$= 7$ ..... (1) 4 Marks.

14. $3xy(2b+q) - 4k(2b+q)$ ..... (1)

$= (3xy-4)(2b+q)$ ..... (2) 3 Marks.

15. $(2x-5)(x+1)$ ..... 3 Marks.

16. $3x+y = 21$ .....(1)
$x+4y = 7$ .....(2)

From (2); $x = 7-4y$ ..... (1)

Subst. in (1)

$3(7-4y) + y = 21$

$21-12y + y = 21$ ..... (1)

$-11y = 21-21$

$-11y = 0$ ....

$y = 0$ ..... (1)

Subst. in (2) $x = 7$ ..... (1) 4 Marks.

(7,0)

Set B: 4

17. $2x-y = 4$ .....(1)
$3x+y = 11$ .....(2)

From (1)

$y = 2x-4$ ..... (1)

Subst. in (2)

$3x+2x-4 = 11$ ..... (1)
$5x-4 = 11$ . . ....
$5x = 11+4$
$= 15$
$x = 15/5$
$= 3$ ..... (1)

Subst. x in (1)

$2x3-y = 4$
$6-y = 4$
$y = 6-4$
$y = 2$ ..... (1) 4 Marks
(3,2) .... ! .

18. $v+ \Pi jy = \overline{\Pi r}^2$
$\Pi jy = \overline{\Pi r}^2 - v$ ..... (1)

Divide through by $\overline{\Pi}j$

i.e. $\frac{\Pi jy}{\overline{\Pi}j} = \frac{\overline{\Pi r}^2 - v}{\overline{\Pi}j}$

$y = \frac{\overline{\Pi r}^2 - v}{\overline{\Pi}j}$ ..... (1) 2 Marks.

19. Let the number be x

Then $\frac{x-7}{3} = \frac{x+5}{2}$ ..... (2)

$\therefore$ $2x-14 = 3x+15$ ..... (1)
$3x-2x = -14-15$ ..... (1)
$x = -29$ .... (1) 5 Marks.

OR

Set B: 5.

Let the number be x.

Then $\frac{x-7}{3} = \frac{x+5}{2}$ ..... (2)

$\therefore \frac{x-7}{3} - \frac{x+5}{2} = 0$

$\frac{2x-14-3x-15}{6}$ ..... (1)

$\frac{2x-14-3x-15}{6} = 0$

$2x-14-3x-15 = 0$ ..... (1)

$-x-29 = 0$
$-x = 29$
$x = -29$ ..... (1) 5 Marks.

22. Let the two numbers be x and y.

Then $x+y = 19$ .....(1) ..... (1)
and $x-y = 5$ .....(2) ..... (1)

(1)+(2) $2x = 24$
$x = 12$ ..... (1)

Subst. x in (1)

$12+y = 19$
$y = 19-12$
$= 7$ ..... (1)

The two numbers are 12 & 7 ..... (1) 5 Marks.

OR By substitution 5 Mks.

OR Graphically 5 Mks.

23. $U = kyz$

Subst. the values of U, y, and Z.

i.e. $64 = k. 4. 8$ ..... (1)
$32k = 64$
$k = \frac{64}{32}$
$= 2$ ..... (1)

$U = k yz$
$= 2.10.5$
$= 100$ ..... (1) 3 Marks.

Set B: 6.

24. $V = \pi r^2 h$

$h = \frac{V}{\pi r^2}$ ..... (1)

$h = \frac{99}{\frac{22}{7}.3^2}$

$= \frac{\overset{9}{\cancel{99}}}{\underset{1}{\cancel{9}}} \times \frac{7}{\underset{2}{\cancel{22}}}$ ..... (1)

$= 7/2$

$= 3\frac{1}{2}$ ..... (1) 3 Marks.

27. More time less students.
Less time more students.

Time $\propto \frac{1}{\text{students}}$ ..... (1)

$\therefore$ T $\propto \frac{1}{S}$ ..... (2) } (3)

$\therefore$ $T = \frac{k}{S}$ ..... (1)

i.e. $7 = \frac{k}{12}$ ..... (1)

$k = 84$

$\therefore$ $T = \frac{\overset{6}{\cancel{84}}}{\underset{1}{\cancel{14}}}$ ..... (1)

$= 6$ days ..... (1)

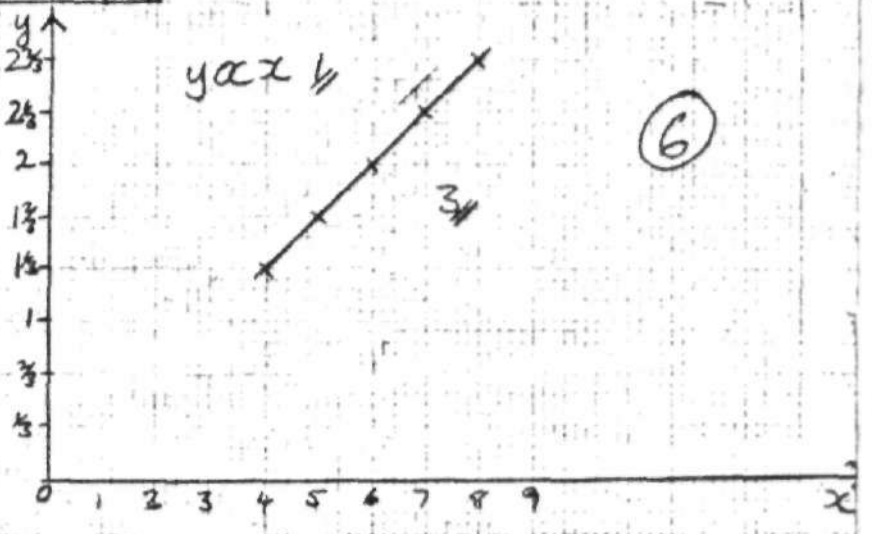
$\therefore$ The bag can feed 14 students
for 6 days ..... (1) 8 Marks.

SETB

19

x	4	.5	6	7	8
y	$1\frac{1}{3}$	$1\frac{2}{3}$	2	$2\frac{1}{3}$	$2\frac{2}{3}$

2



$y \propto x$ ½

3½

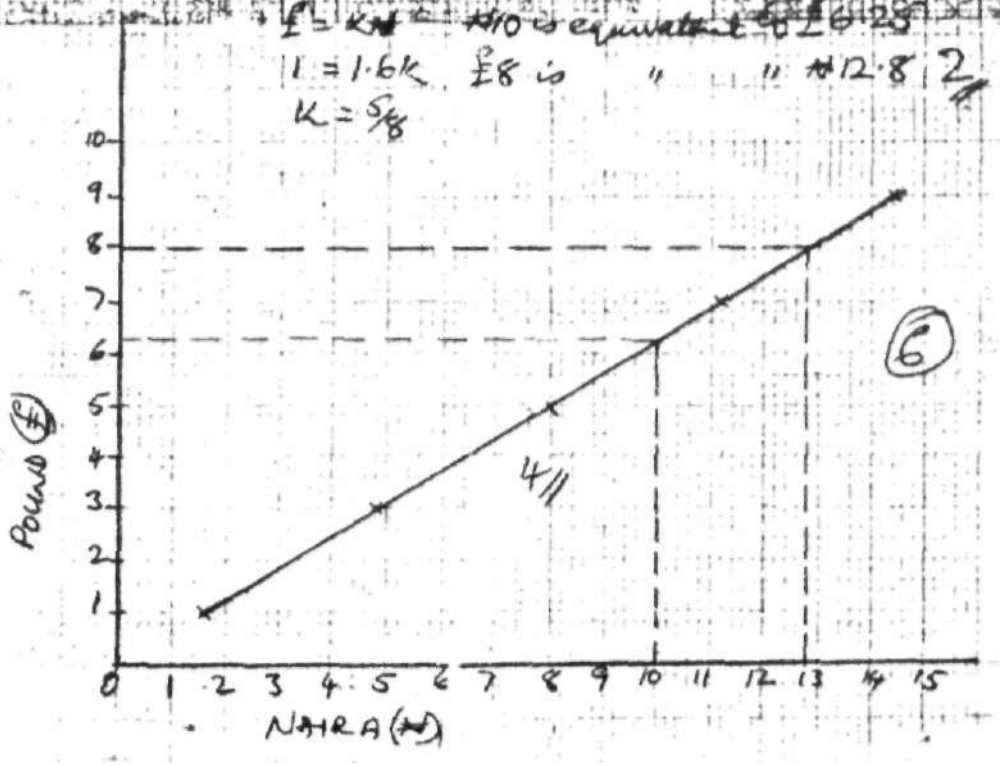
(6)

21

£1 = ₦10 is equivalent to £0.625

1 = 1.6K £8 is " " ₦12.8 2

K = 5/8



4½

(6)

SET B

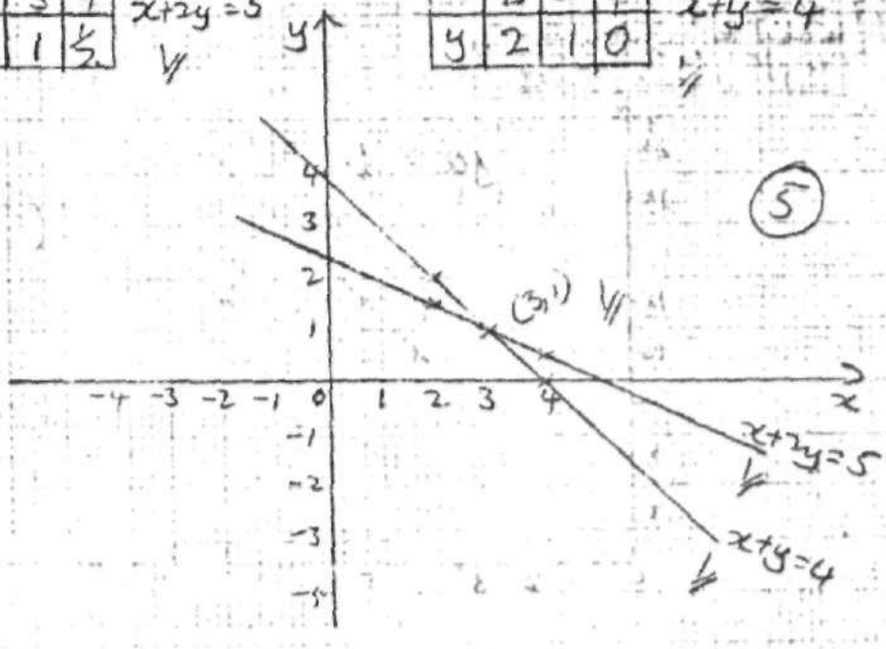
25

x	2	3	4
y	1½	1	½

$x+2y=5$

x	2	3	4
y	2	1	0

$x+y=4$



(5)

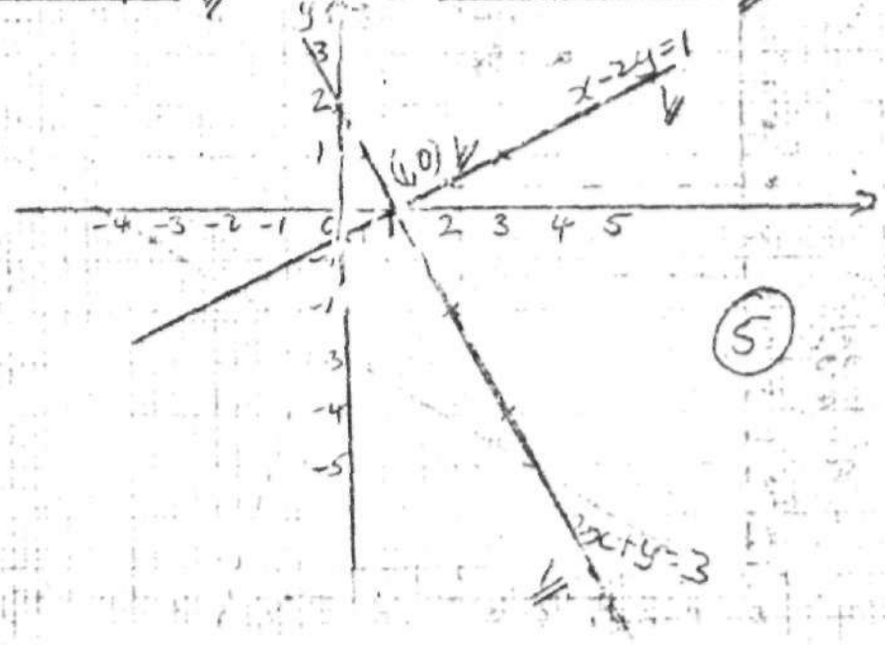
26

x	0	1	2	3
y	2	0	-2	-4

$2x+y=2$

x	0	1	2	3
y	-½	0	½	1

$x-2y=1$



(5)

APPENDIX H

FREQUENCY TABLES OF STUDY DATA

Table 6.2. Frequency Table Of Set A, Total.

CLASS-INTERVAL	FREQUENCY
65-69	3
60-64	1
55-59	1
50-54	1
45-49	6
40-44	8
35-39	11
30-34	28
25-29	42
20-24	56
15-19	115
10-14	153
5-9	234
0-4	697
	1356

Table 6.3. Frequency Table Of Set B. Total.

CLASS-INTERVAL	FREQUENCY
75-79	1
70-74	0
65-69	0
60-64	0
55-59	1
50-54	3
45-49	3
40-44	8
35-39	11
30-34	18
25-29	33
20-24	47
15-19	85
10-14	108
5-9	194
0-4	788
	1300

Table 6.5. Frequcney Table of Set B. Boys.

CLASS-INTERVAL	FREQUCNEY
55-59	1
50-54	1
45-49	1
40-44	2
35-39	6
30-34	6
25-29	11
20-24	17
15-19	20
10-14	46
5-9	92
0-4	575
	778

Table 6.5. Frequency Table of Set A. Girls.

CLASS-INTERVAL	FREQUENCY
50-54	1
45-49	2
40-45	2
35-39	1
30-34	5
25-29	6
20-24	15
15-19	27
10-15	37
5-9	66
0-4	157
	319

Table 6.7. Frequency Table Of Set B. Girls.

CLASS_INTERVAL	FREQUENCY
75-79	1
70-74	0
65-69	0
60-64	0
55-59	0
50-54	2
45-49	2
40-44	2
35-39	0
30-34	4
25-29	6
20-24	16
15-19	30
10-14	35
5-9	70
0-4	187
	355

Table 6.8. Frequency Table Of Set A. Private.

CLASS-INTERVAL	FREQUENCY
65-69	3
60-64	1
55-59	1
50-54	1
45-49	2
40-44	7
35-39	8
30-34	18
25-29	33
20-24	32
15-19	62
10-14	70
5-9	78
0-4	102
	418

Table 6.9. Frequency Table Of Set B, Private.

CLASS-INTERVAL	FREQUENCY
75-79	1
70-74	0
65-69	0
60-64	0
55-59	1
50-54	1
45-49	3
40-44	6
35-39	7
30-34	12
25-29	19
20-24	29
15-19	52
10-14	50
5-9	55
0-4	78
	314

Table 6.17. Frequency Table Of Set A. State Government.

CLASS-INTERVAL	FREQUENCY
45-49	4
40-44	1
35-39	3
30-34	10
25-29	9
20-24	23
15-19	53
10-14	83
5-9	156
0-4	596
	938

Table 6.14 Frequency Table Of Set B, State Government.

CLASS-INTERVAL	FREQUENCY
50-54	2
45-49	0
40-44	2
35-39	4
30-34	6
25-29	14
20-24	18
15-19	33
10-14	58
5-9	139
0-4	710
	986

APPENDIX I

PERCENTILE RANK OF STUDY DATA

Table 6.12 Percentile Rank Set A. Total.

CLASS-INTERVAL	F	CF	CF_{mp}	CP	PR
65-69	3	1353	1354.5	.999	99.9
60-64	1	1352	1352.5	.997	99.7
55-59	1	1351	1351.5	.996	99.6
50-54	1	1350	1350.5	.995	99.5
45-49	6	1344	1347	.993	99.3
40-44	8	1336	1340	.988	98.8
35-39	11	1325	1330.5	.981	98.1
30-34	28	1297	1311	.967	96.7
25-29	42	1255	1276	.941	94.1
20-24	56	1199	1227	.905	90.5
15-19	115	1084	1141.5	.842	84.2
10-14	153	931	1007.5	.743	74.3
5-9	234	697	814	.600	60.0
0-4	697	0	348.5	.257	25.7
	1356				

Table 6.13 Percentile Rank Set B. Total

CLASS-INTERVAL	F	CF	CF_{up}	CP	PR
75-79	1	1299	1299	.999	99.9
70-74	0	1299	1298.5	.998	99.8
65-69	0	1299	1298.5	.998	99.8
60-64	0	1299	1298.5	.998	99.8
55-59	1	1298	1298.5	998	99.8
50-54	3	1295	1296.5	.997	99.7
45-49	3	1292	1293.5	.995	99.5
40-44	8	1284	1288	.991	99.1
35-39	11	1273	1278.5	.983	98.3
30-34	18	1255	1264.	.972	97.2
25-29	33	1222	1238.5	.953	95.3
20-24	47	1175	1198.5	.922	92.2
15-19	85	1090	1132.5	.871	87.1
10-14	108	982	1036	.797	79.7
5-9	194	788	885	.681	68.1
0-4	788	0	364	.280	28.0
	1300				

Table 6.14 Percentile Rank Set A. Boys.

CLASS-INTERVAL	F	CF	CF_{mp}	CP	PR
65-69	2	836	837	.999	99.9
60-64	0	836	835.5	.997	99.7
55-59	1	835	835.5	.997	99.7
50-54	1	834	834.5	.996	99.6
45-49	2	832	833	.994	99.4
40-44	4	828	830	.990	99.0
35-39	5	823	825.5	.985	98.5
30-34	13	810	816.5	.974	97.4
25-29	15	795	802.5	.958	95.8
20-24	23	772	783.5	.934	93.4
15-19	55	717	744.5	.888	88.8
10-14	81	636	676.5	.807	80.7
5-9	131	505	570.5	.681	68.1
0-4	505	0	252.5	.301	30.1
	838				

Table 6.15 Percentile Rank Set B. Boys.

CLASS INTERVAL	F	CF	CF_{pm}	CP	PR
55-59	1	777	777.5	.999	99.9
50-54	1	776	776.5	.998	99.8
45-49	1	775	775.5	.997	99.7
40-44	2	773	774	.995	99.5
35-39	6	767	770	.990	99.0
30-34	6	761	764	.982	98.2
25-29	11	750	755.5	.971	97.1
20-24	17	733	741.5	.953	95.3
15-19	20	713	723	.930	93.0
10-14	46	667	700	.900	90.0
5-9	92	575	621	.798	79.8
0-4	575	0	287.5	.370	37.0
	778				

Table 6.16 Percentile Rank Of Set A. Girls.

CLASS-INTERVAL	F	CF	CF_{np}	CP	PR
50-54	1	318	318.5	.998	99.8
45-49	2	316	317	.994	99.4
40-44	2	314	315	.987	98.7
35-39	1	313	313.5	.983	98.3
30-34	5	308	310.5	.973	97.3
25-29	6	302	305	.956	95.6
20-24	15	287	294.5	.923	92.3
15-19	27	260	273.5	.857	85.7
10-14	37	223	241.5	.757	75.7
5-9	66	157	190	.596	59.6
0-4	157	0	78.5	.246	24.6
	319				

Table 6.17 Percentile Rank Of Set B. Girls.

CLASS-INTERVAL	F	CF	CF_{mp}	CP	PR
75-79	1	354	354.5	.999	99.9
70-74	0	354	354	.997	99.7
65-69	0	354	354	.997	99.7
60-64	0	354	354	.997	99.7
55-59	0	354	354	.997	99.7
50-54	2	352	353	.994	99.4
45-49	2	350	351	.989	98.9
40-44	2	348	349	.983	98.3
35-39	0	348	348	.980	98.0
30-34	4	344	346	.975	97.5
25-29	6	338	341	.961	96.1
20-24	16	322	330	.930	93.0
15-19	30	292	307	.865	86.5
10-14	35	257	274.5	.773	77.3
5-9	70	187	222	.625	62.5
0-4	187	0	93.5	.263	26.3
	355				

Table 6.18 Percentile Rank of Set A. Private.

CLASS INTERVAL	F	CF	CP mp	CP	PR
65-69	3	415	416.5	.996	99.6
60-64	1	414	414.5	.992	99.2
55-59	1	413	413.5	.989	98.9
50-54	1	412	412.5	987	98.7
45-49	2	410	411	.983	98.3
40-44	7	403	406.5	972	97.2
35-39	8	395	399	.955	95.5
30-34	18	377	386	.923	92.3
25-29	33	344	360.5	.862	86.2
20-24	32	312	328	.785	78.5
15-19	62	250	281	.672	67.2
10-14	70	180	215	.514	51.4
5-9	78	102	138	.330	33.0
0-4	102	0	51	.122	12.2
	418				

Table 6.19 Percentile Rank Of Set B. Private.

CLASS-INTERVAL	F	CF	CF_{mp}	CP	PR
75-79	1	313	313.5	.998	99.8
70-74	0	313	313	.997	99.7
65-69	0	313	313	.997	99.7
60-64	0	313	313	.997	99.7
55-59	1	312	312.5	.995	99.5
50-54	1	311	311.5	.992	99.2
45-49	3	308	309.5	.986	98.6
40-44	6	302	305	.971	97.1
35-39	7	295	298.5	.951	95.1
30-34	12	283	289	.920	92.0
25-29	19	264	273.5	.871	87.1
20-24	29	235	249.5	.795	79.5
15-19	52	183	209	.666	66.6
10-14	50	133	155	.494	49.4
5-9	55	78	105.5	.336	33.6
0-4	78	0	39	.124	12.4
	314				

Table 6.20 Percentile Rank of Set A. State Government.

CLASS INTERVAL	F	CF	CF_{mp}	CP	PR
45-49	4	934	936	.998	99.8
40-44	1	933	933.5	.995	99.5
35-39	3	930	931.5	.993	99.3
30-34	10	920	925	.986	98.6
25-29	9	911	915.5	.976	97.6
20-24	23	888	899.5	.959	95.9
15-19	53	835	861.5	.918	91.8
10-14	83	752	793.5	.846	84.6
5-9	156	596	674	.719	71.9
0-4	596	0	298	.318	31.8
	938				

Table 6.21 Percentile Rank Of Set B.
State Government.

CLASS-INTERVAL	F	CF	CF_{up}	CP	PR
50-54	2	984	985	.999	99.9
45-49	0	984	984	.998	99.8
40-44	2	982	983	.997	99.7
35-39	4	978	980	.994	99.4
30-34	6	972	975	.989	98.9
25-29	14	958	965	.979	97.9
20-24	18	940	949	.962	96.2
15-19	33	907	913.5	.926	92.6
10-14	58	849	878	.890	89.0
5-9	139	710	779.5	.791	79.1
0-4	710	0	355	.360	39.0
	986				

APPENDIX J

NORMALIZED STANDARD SCORE OF STUDY DATA

Table 6.22 Normalized Standard Scores Of Set A. Total

CLASS INTERVAL	CP	Z	T
65-69	.999	3.0	80
60-64	.997	2.8	78
55-59	.996	2.7	77
50-54	.995	2.6	76
45-49	.993	2.5	75
40-44	.988	2.3	73
35-39	.981	2.1	71
30-34	.967	1.8	68
25-29	.941	1.6	66
20-24	.905	1.3	63
15-19	.842	1.0	60
10-14	.743	0.7	57
5-9	.600	0.3	53
0-4	.257	-0.7	43

Tabl 6.23 Normalized Standard Scores Of: Set B. Total

CLASS-INTERVAL	CP	Z	T
75-79	.999	3.0	80
70-74	.998	2.9	79
65-69	.998	2.9	79
60-64	.998	2.9	79
55-59	.998	2.9	79
50-54	.997	2.8	78
45-49	.995	2.6	76
40-44	.991	2.4	74
35-39	.983	2.1	71
30-34	.972	1.9	69
25-29	.953	1.7	67
20-24	.922	1.4	64
15-19	.871	1.1	61
10-14	.797	0.8	58
5-9	.681	0.5	55
0-4	.200	-0.6	44

Table 6.24 Normalized Standard Scores Of Set A. Boys.

CLASS-INTERVAL	CP	Z	T
65-69	.999	3.0	80
60-64	.997	2.8	78
55-59	.997	2.8	78
50-54	.996	2.7	77
45-49	.994	2.5	75
40-44	.990	2.3	73
35-39	.985	2.2	72
30-34	.974	1.9	69
25-29	.958	1.7	67
20-24	.934	1.5	65
15-19	.888	1.2	62
10-14	.807	0.9	59
5-9	.681	0.5	55
0-4	.301	-0.5	45

Table 6.25 Normalized Standard Scores Of Set B. Boys.

CLASS-INTERVAL	CP	Z	T
55-59	.999	3.0	80
50-54	.998	2.9	79
45-49	.997	2.8	78
40-44	.995	2.6	76
35-39	.990	2.3	73
30-34	.982	2.1	71
25-29	.971	1.9	69
20-24	.953	1.7	67
15-19	.930	1.5	65
10-14	.900	1.3	63
5-9	.798	0.8	58
0-4	.370	-0.3	47

Table 6.26 Normalized Standard Scores Of Set A. Girls.

CLASS-INTERVAL	CP	Z	T
50-54	.998	2.9	79
45-49	.994	2.5	75
40-44	.987	2.2	72
35-39	.983	2.1	71
30-34	.973	1.9	69
25-29	.956	1.7	67
20-24	.923	1.4	64
15-19	.857	1.1	61
10-14	.757	0.7	57
5-9	.596	0.2	52
0-4	.246	-0.7	43

Table 6.27 Normalized Standard Scores Of Set B. Girls.

CLASS-INTERVAL	CP	Z	T
75-79	.999	3.0	80
70-74	.997	2.8	78
65-69	.997	2.8	78
60-64	.997	2.8	78
55-59	.997	2.8	78
50-54	.994	2.5	75
45-49	.989	2.3	73
40-44	.983	2.1	71
35-39	.980	2.1	71
30-34	.975	2.0	70
25-29	.961	1.8	68
20-24	.930	1.5	65
15-19	.865	1.1	61
10-14	.773	0.8	58
5-9	.625	0.3	53
0-4	.263	-0.6	44

Table 6.28 Normalized Standard Scores Of Set A. Private.

CLASS-INTERVAL	CP	Z	T
65-69	.996	2.7	77
60-64	.992	2.4	74
55-59	.989	2.3	73
50-54	.987	2.2	72
45-49	.983	2.1	71
40-44	.972	1.9	69
35-39	.955	1.7	67
30-34	.923	1.4	64
25-29	.862	1.1	61
20-24	.785	0.8	58
15-19	.672	0.5	55
10-14	.514	0.0	50
5-9	.330	-0.4	46
0-4	.122	-1.2	38

Table 6.29 Normalized Standard Scores Of Set B. Private.

CLASS-INTERVAL	CP	Z	T
75-79	.998	2.9	79
70-74	.997	2.8	78
65-69	.997	2.8	78
60-64	.997	2.8	78
55-59	.995	2.6	76
50-54	.992	2.4	74
45-49	.986	2.2	72
40-44	.971	1.9	69
35-39	.951	1.7	67
30-34	.920	1.4	64
25-29	.871	1.1	61
20-24	.795	0.8	58
15-19	.666	0.4	54
10-14	.494	0.0	50
5-9	.336	-0.4	46
0-4	.124	-1.2	38

Table 6.30 Normalized Standard Scores Of Set A. State Government.

CLASS-INTERVAL	CP	Z	T
45-49	.998	2.9	79
40-44	.995	2.6	76
35-39	.993	2.5	75
30-34	.986	2.2	72
25-29	.976	2.0	70
20-24	.959	1.7	67
15-19	.918	1.4	64
10-14	.846	1.0	60
5-9	.719	0.6	56
0-4	.318	-0.5	45

Table 6.31 Normalized Standard Scores Of Set B. State Government.

CLASS-INTERVAL	CP	Z	T
50-54	.999	3.0	80
45-49	.998	2.9	79
40-44	.997	2.8	78
35-39	.994	2.5	75
30-34	.989	2.3	73
25-29	.979	2.0	70
20-24	.962	1.8	68
15-19	.926	1.4	64
10-14	.890	1.2	62
5-9	.791	0.8	58
0-4	.360	-0.4	46