A PSYCHOMETRIC ASSESSMENT OF SCHOLASTIC APTITUDE TESTS USED IN ZAMBIA

BY

GOLIATH BERRY SILUCHALI

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE MASTER OF EDUCATION

UNIVERSITY OF ZAMBIA

1989

DEDICATION

I dedicate this thesis to my father and mother, my son Katanga and my daughters Munsanda, Masandu, Namaloba, and Silimi.

ACKNOWLEDEMENT

I wish to make grateful acknowledgement to the following:

(i) Dr. R.F. Zimba of the University of Zambia, who as my supervisor, made helpful observations and suggestions on my work without which, this dissertation could not have reached its present standard.

(ii) The personnel at the Examinations Council of Zambia for permission to use the Special Paper past examination question booklets.

(iii) The Headmasters of Lake Road, Tree Tops, Mumuni, and Twatasha Schools for allowing me to conduct my research in their schools, and the members of staff for their co-operation.

(iv) My sponsors, the Directorate of Manpower Development and Training, for the financial support.

(v) The Council for the Development of Economic and Social Research in Africa (CODESRIA), for the grant to enable me to meet the extra expenses, and most of all, their ordering on my behalf, the books and journals that I required for my thesis. I should mention here that because my research exceeded the 18 months that our

government sponsorship normally covers, it would have been impossible for me to complete the study.

(vi) Kennedy Zulu of Evelyn Hone College for the cartoon, my mulamu S. Tabali, my sister Maliana, my nephew [illegible], my friends Russel Mushanga, Lazarus Hikunuma, and Spencer Munachonga Simunde for taking the responsibility of printing this thesis.

(vii) Last, but not least, my employers, the Office of the President, for granting me the study leave.

DECLARATION

I, Goliath Berry SILUCHALI, declare that this dissertation has been composed by me, and that the work recorded is my own. All quotations, if any, have been distinguished by quotation marks. The sources of all materials used have been specifically acknowledged, and the dissertation has not been previously submitted for a degree or another academic attainment at this or another University.

Signature G. B. Siluchali

Date: 28th May, 1990

APPROVAL

This dissertation of **Goliath Berry Siluchali** is approved as fulfilling part of the requirements for the award of the degree of Master of Education by the University of Zambia.

Signature:	Date:
.....	11/11/90
.....	26/03/91
.....	26/03/91
.....	.....

ABSTRACT

A PSYCHOMETRIC ASSESSMENT OF SCHOLASTIC APTITUDE TESTS USED IN ZAMBIA

To select secondary school entrants, Zambia uses a special paper as a scholastic aptitude test. The paper has a Western origin and even though it has been in use for more than two decades, it has no manual and its validity and reliability have not been assessed.

The present study's objective was to assess the validity and reliability of the special paper. The study was experimental. Pupils from four schools were selected as subjects. The schools were selected on the basis of their socio-economic status, thus the high, the low and the intermediate socio-economic groups. The subjects were selected on the basis of their performance in school subjects. From each school, two groups; bright pupils, and weak pupils were formed. Thus the high ability and the low ability groups respectively. Two of the schools formed the experimental group whereas the other two, the control group.

The experiment was in two phases; pretest and posttest. Test scores were compared between ability groups, socio-economic groups, and the phases. The paper's subject matter and construction were also assessed.

The results obtained showed that the paper was susceptible to socio-economic background and coaching. It was also found out that the subject matter was of low level and that construction of items was in some cases ambiguous and even erroneous. However, the test indicated good correlation between performance in school subjects and performance on it.

TABLE OF CONTENTS

List of Tables

CHAPTER ONE

BACKKGROUND TO THE PROBLEM AND REVIEW OF RELATED LITERATURE

CHAPTER TWO

THE PROBLEM

CHAPTER THREE

METHODOLOGY

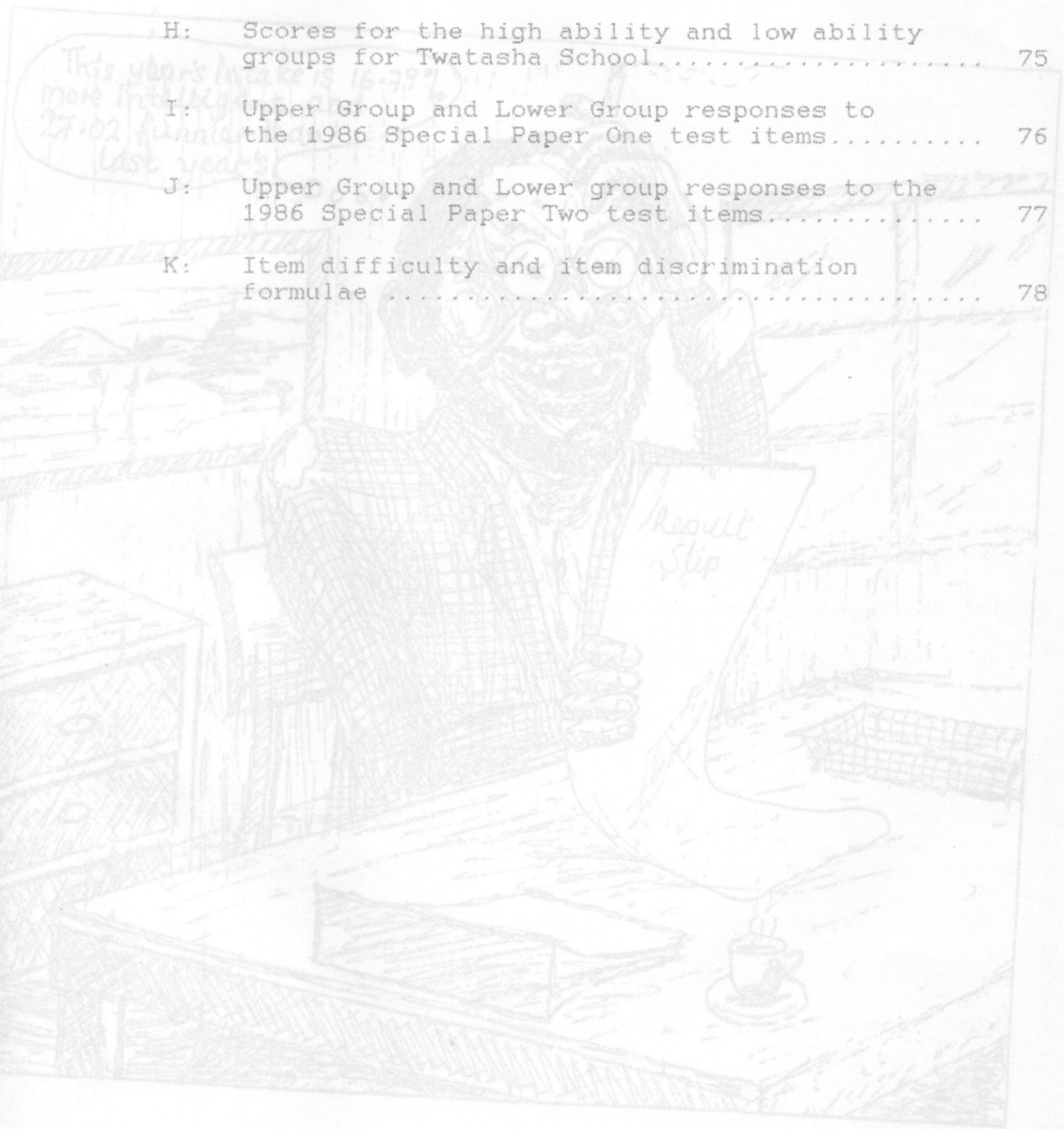
CHAPTER FOUR

RESULTS

CHAPTER FIVE

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

APPENDIX




This year's intake is 16.79% more intelligent and 27.02 funnier than last year's.....
Result Slip

CHAPTER ONE

BACKGROUND TO THE PROBLEM AND REVIEW OF RELATED LITERATURE

BACKGROUND TO THE PROBLEM

Scholastic aptitude, being intellectual, lacks objectives criteria for its measurement. In an attempt to assess it, psychometricians choose a set of tasks aimed at creating a miniature environment within which the potential for learning can be demonstrated (Serpell, 1972). The appropriateness of these tasks, however, needs to be verified. This verification can be done by looking at the validity and reliability of these measures.

In order to select secondary school pupils, those pupils who have the potential for further learning, Zambia uses a scholastic aptitude test. This test is called the Special Paper. The test has two parts; paper one and paper two. Whereas paper one is verball, paper two is non-verbal.

The Special Paper test has a Western origin. Except for

the present decade, the test was constructed by the University of Cambridge. For nearly ten years now, the test has been prepared by the Examinations Council of Zambia. Each year, the council prepares the test items by modifying or reproducing the old items which were constructed by the University of Cambridge.

After the examination, the question papers are withdrawn from the schools. Test items are therefore exposed to pupils at examination times only.

Although the Special Paper test has been in use for more than two decades, it has never been seriously reviewed or assessed for its validity and reliability. In fact, there is no test manual for the paper in Zambia. This means that very little documented evidence on the test's validity and reliability in the Zambian educational and cultural setting exists.

REVIEW OF RELATED LITERATURE

In cross cultural research the concern with intelligence tests has been comparability of test scores. Thus, whether the same quantitative and qualitative meaning can be attached to test scores in different cultural groups.

Vernon's conception of intelligence (Vernon 1955) is of relevance to cross cultural research. In his study, Vernon identifies three levels of intelligence; Intelligence A, B and C. He regards Intelligence A as being an inborn characteristic, and Intelligence B as being the product of interaction between the genetic potential of those kinds of behaviour which are called Intelligent, but does not, of itself, imply an entity. Intelligence C is what is measured by a standardised test; it is a score on an intelligence test.

Comparability in respect of Vernon's conception of intelligence is concerned with the relationship between Intelligence B and Intelligence C, thus the influence of environmental or cultural factors on performance on an intelligence test. The relationship which the three levels of intelligence have, has raised expectations that if the effects of Intelligence B are neutralised, Intelligence C will be the measure of Intelligence A applicable to all cultural groups. This expectation has led to the construction of culture free tests, requiring behaviour display of only genetic characteristics.

Substantial study in cross cultural research has been done on culture free tests, and Vernon has been one of the contributors. His study (1969) can be cited on this subject. The study suggests that each culture actively

promotes the learning of certain skills that are relevant to the performance of tasks which are considered important. Thus, in cultures where literacy and technology have become of paramount importance in education, a lot of skills concerned with deciphering symbols on bits of paper have come to be seen as essential and are taught to almost all children at a very early age. In many of the rural communities in Africa, on the other hand, such skills are irrelevant to the valued activities of the local culture and are therefore not taught unless children enter the exceptional environment of formal schooling. Even then, they receive far less emphasis than in countries such as Britain and the U.S.A.

The importance of Vernon's contribution is supported by Deregowski and Serpell's 1971 study in Lusaka. In this study, African children scored rather low on tests that required copying of shapes or designs with pencil and paper. On the other hand, Western children scored rather low on tests that required wire bending, for instance, to build skeleton cars out of old bits of wire. Evidently, Western children out scored their African counterparts on the former tests and African children outscored their Western counterparts on the latter tests because drawing with pencil is a skill which Western children pick up at home and in school

very early in life, and making car models out of wire is a skill which children growing up in a Zambian city pick up.

The effect of cultural setting on performance on a particular test, can be further supported by the research carried out by a group of workers at the University of London in the late 1920's. The research was on the similarities and differences between intelligence tests based on perceptual material and those of determined material. Subsequently, a number of perceptual tests for practical use, such as Pintner, Porteus, and Reven were used. The subject was taken up systematically again in the late 1930's by Cattel, Sarason, and others who experimented both with the perceptual material and with the generally familiar material, such as parts of the body and geographically common features of the environment. In this study, it was shown that an immigrant group in the United States tested on traditional tests, the Binet and the American College Entrance Examination (ACE) along the Culture Free Tests, improved on a second testing only on the Binet and the ACE. Thus both the perceptual tests and the performance tests showed to have some advantages over the traditional type of tests in that they were less susceptible to cultural influences.

Although Culture Free Tests have shown little susceptibility to cultural influences, we can not unquestionably justify their use in all settings because the potential to tackle a test, which in this case we shall refer to as intelligence, is complex and as such, it's assessment may require more than simply a Culture Free Test.

An intelligente test's validity and reliability may therefore need to be assessed to determine whether or not the test tapes the traits necessary for the exhibition of the required behaviour and the extent to which it can produce the desirable result.

A rather limited review of some aspects of the Special Paper's validity was conducted by the Psychological Services of Zambia in 1973. The purpose of the review was to determine whether or not those pupils who scored highly on the special paper test would also score highly in technical subjects. This study was worthwhile because by then (1973) entrants into technical secondary schools were selected on the basis of their performance on the Special Paper test. The study included only David Kaunda and Hillcrest Technical Secondary Schools. In this study, no correlation was found between performance on Special Paper and performance in Metalwork, Woodwork, Mathematics, Physics, Chemistry,

and Biology. However, a correlation of 0.53 was found between Special Paper Two and Technical drawing.

Encouraged by these results, the Psychological Service was to, in 1974, embark on a study to determine the special skills that Special Paper Two tapped so that it could accordingly advise the Ministry of Education to treat Special Paper One and Special Paper Two as separate tests. However, the results of this study were not published, and the service has not published any Annual Report since 1973. The two papers, as a result, are still treated as one test.

CHAPTER TWO

THE PROBLEM

STATEMENT OF THE PROBLEM

The present study narrows Vernon's 1969 study in terms of cross-cultural validity, and broadens the Psychological Service's 1973 study and Vernon's 1955 study in terms of predictive and construct validity. In addition, it looks at the paper's reliability. Thus, the study attempts to assess the psychometric quality of the special paper in a more complete manner. This assessment tries to answer the following questions:-

(1) Does the special paper test measure scholastic aptitude which it is intended to measure?

(2) Does the test yield consistent results over a period of time?

The first question refers to validity, and the second one to reliability of the test. To determine validity, the study examines factors that are likely to affect the effectiveness of a test in serving the purpose which it is intended for. To do this, answers are sought to the question of whether or not the subject matter and the mental functions that are necessary for school

achievement are measured by the test. In a scholastic aptitude test, test items should among other traits measure: (i) the ability of the pupil to translate and transfer information from the area of origin to other areas, (ii) the ability of the pupil to break down, compare and contrast information, (iii) the ability of the pupil to combine ideas into new ones, and (iv) the ability of the pupil to make judgements about the values of the information. A test which attempts to measure these complex skills with items which are appropriate only for measuring factual knowledge is less valid as a scholastic aptitude test.

The present study looks at the special paper's level of difficulty, construction, and sentence structure to assess the absence or presence of ambiquity in its items.

The level of difficulty of a test refers to the proportion of pupils who on average, are able to answer its questions correctly. The level of difficulty of an aptitude test, for instance the special paper, which is supposed to measure the pupils' ability to deal with propositions, to imagine possibilities, to formulate and test hypotheses should be matched with those pupils' intellectual understanding. According to Piaget (1947), it is during the formal operations stage of cognitive

development that pupils display the ability to handle propositions. Although the onset of this period may be delayed, it normally becomes available to children by the age of 12 years. Piaget illustrates the absence of the period's mode of thinking by the following example:

> Edith is fairer than Susan; Edith is darker than Lilly; who is the darkest of the three?

Piaget explains that when children who are less than 12 years of age are presented with the above problem, they reason erroneously as follows:

> Edith and Susan are fair; Edith and Lilly are dark; therefore Lilly is the darkest.

To under 12-year-olds, this problem would thus be found to be very difficult. A scholastic aptitude test which assesses skills involving propositions should be administered to children of not less than twelve years of age. Does the special paper meet this requirement?

Construct validity is another important aspect in test development. If test items are poorly constructed, they may unintentionally provide clues to answers and the test will as a result measure the pupils alertness in

detecting clues and not only the ability the test is intended to measure. Similarly, if the test maker uses unfamiliar vocabulary, the test may result in measuring comprehension and not some intended traits.

The arrangement of test items ideally should be in order of difficulty, with easy items placed first. If difficult items are places first, they may cause the pupils to spend too much time on them (difficult items) and prevent them from reaching the items which they can easily answer. Such an arrangement may therefore reduce the validity of a test.

Test items should also be stated clearly. Ambiguous statements contribute to misintepretation and confusion. Test taking directions should be clear too. Directions which do not clearly indicate to the pupil how he should respond to the items tend to reduce the validity of the test.

Length is also an important aspect in test validity. Since a test is only a sample of the many questions that might be asked, it should be long enough to allow the test taker to demonstrate the appropriate amount of the trait the test is supposed to measure.

If the above qualities are taken into account when

constructing a test, we would expect the test to yield a high validity. As alluded to earlier, the present study was aimed at assessing the validity of the special paper test through examining the reviewed aspects. It was hoped that this assessment would provide an answer to the first question: Does the Special Paper Test measure scholastic aptitude which it is intended to measure? Does it yield results which indicate that high ability pupils perform better on it than low ability pupils? In other words, does the Special Paper discriminate (positively) between high ability and low ability pupils?

In addition, a test of scholastic aptitude should predict future scholastic performance. To assess this we need evidence indicating that individuals who obtain high scores on the Special Paper test (at primary school) also obtain high scores in school subjects at secondary school. This information would help us to answer the question of whether or not the test has predictive validity. Another important aspect would be to find out whether performance on the special paper test would be influenced by socio-economic background.

HYPOTHESES TESTED

In view of the preceding discussion, the following hypotheses were constructed:

(1) The higher a pupil's performance in primary school subjects, the higher will be his score on the special paper test.

(2) The higher a pupil's performance on the special paper test, the higher will be his performance in school subjects at secondary school.

(3) The higher a group's socio-economic status, the higher will be the performance of its members on the special paper test.

(4) Performance on the special paper test improves with direct teaching.

CHAPTER THREE

METHODOLOY

SUBJECTS

The subjects were drawn from a population of Grade 6 primary school pupils. The age at which children start primary school in the Zambian education system is 6 or 7 years, so the subjects were 12 or 13 years old.

Two government and two private schools were the sources of the subjects. Whereas Lake Road and Tree Tops were private schools, Mumuni and Twatasha were government schools. From each of these schools, 20 pupils were selected on the basis of their performance in school work. 10 out of the 20 pupils were those rated 'bright' by the teachers, and the other 10 were those rated 'weak'. The bright pupils composed the low ability group. This selection was aimed at making it possible to test the hypothesis that the higher a pupil's performance in primary school subjects, the higher will be his score on the special paper test. It also enabled the researcher to determine the difficulty level of the test items.

The schools were selected on the basis of their socio-economic status. This was done with a view to testing the hypothesis that the higher a group's socio-economic status, the higher will be the performance of its members on the special paper test. The details of the sample schools are as follows:

Lake Road is situated in a sparsely populated area in the outskirts of Lusaka which is inhabited by persons from the high socio-economic status group. It is private, fee paying and comprises both primary and secondary classes. It is attended by pupils from the high socio=economic status group.

Tree Tops is also situated in a sparsely populated area of Lusaka. It is private, fee paying and comprises only primary school classes.

Mumuni is government and non-fee paying. It is situated in a moderately populated area of Lusaka which is inhabited by civil servants.

Twatasha is also a government and non-fee paying school. It is situated in a densely populated area of Lusaka which is inhabited by lower civil servants and the unemployed.

Tree Tops and Twatasha formed the treatment group whereas Lake Road and Mumuni formed the control group.

RESEARCH MATERIALS

The 1986 and 1987 examination papers for the Special Paper One and Special Paper Two tests were used. Each subject was provided with an examination booklet and an answer sheet. Pencils and erasers were provided by the subjects themselves.

Each of the examination papers comprised 50 multiple choice questions and was allowed a duration of 60 minutes. The examination booklets were provided by the Examinations Council and reproduced by the Government Printer on request from the University of Zambia.

PROCEDURE

On production of an introductory letter from the University of Zambia, permission was granted by the Headmasters of the schools to conduct the research. The experiment was done in two phases. Phase one was the pretest whereas phase two was the posttest. The 1986

paper was used for the posttest. The use of different years' papers for the pretest and posttest was possible because the special paper test items are selected from the same pool of items each year. Thus a comparison of performance on any years' papers would still be valid. This means that the procedure was still in a position to yield 'alternate form' type of reliability.

For both phases, both Special Paper One and Special Paper Two were administered. Paper One was written in the morning whereas Paper Two was written in the afternoon of each test day. This procedure was similar to the one that obtains when this examination is taken at Grade Seven.

Phase one of the experiment was conducted in the first week of April, 1989 and phase two in the second week of July, 1989. All the tests were administered by the researcher.

During the period "1st June - 7th July 1989" the treatment group received direct teaching. This meant that the researcher, with pupils actively participating, went through the test items of the pretest paper, outlining clues that were necessary for one to answer the items correctly. Pupils from treatment group

schools received a total of six hours of direct teaching.

DATA ANALYSIS

To analyse the data, the following tasks were carried out:-

(i) The 1986 paper was scrutinised for the items' suitability to measure the outcome of intellect, item difficulty, discrimination power, sentence construction, and clarity. This was necessary because these are the factors most likely to reduce the validity of a test if they are not taken into account.

(ii) Pretest scores for the high ability group and the low ability group pupils were compared statistically using a t-test. This was necessary to test the hypothesis that the higher a pupil's performance in primary school subjects, the higher will be his score on the special paper test. The use of a t-test was appropriate because the samples were related in that the same subjects' performance in the pretest and posttest were

compared with a view to determining the effect of the treatment, thus the influence of coaching on performance on the test.

(iii) Pretest and posttest scores for the treatment group schools were compared, and the same was done for the control group schools. This was necessary to test the hypothesis that performance on the special paper test improves with direct teaching. It was necessary also, to assess reliability.

(iv) Pretest scores were analysed according to socio-economic status. This was necessary to test the hypothesis that the higher a group's socio-economic status, the higher will be the performance of its members on the special paper test.

(v) The 1985 special paper scores for the highest 10 and the lowest 10, and the 1987 Junior Secondary School Leaving Examination (JSSLE) scores for Lake Road pupils were correlated. This was done in order to test the hypothesis that the higher a pupil's performance on the special paper test, the higher will be his performance in school subjects at secondary school.

(iv) Discrimination power of the test was worked out. This was done to determine the extent to which the test distinguished high ability pupils from low ability ones. Knowing this was judged essential in assessing the special paper as a selection test.

CHAPTER FOUR

RESULTS

The comparison of pretest scores for the high ability and the low ability groups, which was testing the hypothesis that the higher a pupil's performance in school subjects, the higher will be his score on the special paper test (SPT), yielded results as shown in table 1 below.

Table 1: High ability and low ability groups scores compared.

SCHOOL	HIGH ABILITY GROUP			LOW ABILITY GROUP			t VALUE	CRITICAL VALUE for = 0.5
	NO. of Subjects	S P mean score %	SPSD score	NO. of subjects	S P mean score %	S P S.D. score		
Lake Road	10	74	7.71	10	69	6.02	1.358	2.262
Mumuni	10	73	7.42	10	36	37.6	3.054	2.262
Tree Tops	10	83	4.47	10	67	10.42	4.518	2.262
Twatasha	10	53	13.4	10	28	17.76	3.653	2.262

Mumuni, Tree Tops and Twatasha had the high ability groups scoring significantly higher than the low ability groups; $t(9) = 3.054$, $P < .05$, $t(9) = 4.518$, $P < 0.5$, and $t(9) = 3.653$, $P < .05$ respectively.

Lake Road also had the high ability group scoring higher than the low ability. However, this difference was not significant at 0.05 level.

The comparison of the 1985 special paper scores of the top 10 and bottom 10 pupils from Lake Road school with their 1987 JSSLE performance which was aimed at testing the hypothesis that the higher a pupil's performance on the special paper test, the higher will be his performance in school subjects at secondary school (Grade 9) yielded the results given in table 2.

Table 2. Rank order correlation coefficient for the 1985 special paper and the 1987 JSSLE scores for the top 10 and bottom 10 pupils from Lake Road school.

PUPIL	SPECIAL PAPER SCORE	JSSLE SCORE	RANK IN SPECIAL PAPER	RANK IN JSSLE	DIFFERENCE IN RANK (D)	DIFFE-RENCE SQUIRED (D) 2
A	85	85	1	4	-3	9
B	84	86	2	2	0	0
C	82	84	3	6	-3	9
D	81	86	4	2	2	4
E	78	88	5	1	4	16
F	78	84	6	6	0	0
G	78	72	7	13	-6	36
H	76	84	8	6	2	4
I	76	83	9	9	0	0
J	74	80	10	10	0	0
K	66	75	11	11	0	0
L	65	58	12	20	-8	64
M	64	85	13	4	9	81
N	64	75	14	11	3	9
O	61	61	15	17	-2	4
P	55	68	16	14	2	4
Q	54	67	17	15	2	4
R	53	67	18	15	3	9
S	46	55	19	19	0	0
T	45	59	20	18	2	4
						D=257

The special paper and the JSSLE scores yielded a rank-order correlation of 0.8.

When pretest scores were compared amongst the sampled schools in testing the hypothesis that the higher a group's socio-economic status, the higher will be its performance on the special paper test, the results were as is shown in table 3 below:-

Table 3: Pretest scores analysed according to socio-economic groups.

REFERENT SCHOOL	NO. OF SUBJECTS	SPT SCORE MEAN	SPT SCORE S.D	SCHOOL COMPARED WITH	t VALUE	CRITICAL VALUE FOR = 0.05
LAKE ROAD	20	71	7.18	Twatasha Mumuni	6.425 3.969	2.093 2.093
MUMUNI	20	54	17.76	Tree Tops	4.450	2.093
TREE TOPS	20	75	11.40	Lake Road	1.311	2.093
TWATASHA	20	40	20.20	Mumuni Tree Tops	2.328 6.700	2.093 2.093

Lake Road, a high socio-economic status (HSES) school scored significantly higher than Twatasha and Mumuni which were both low socio-economic status (LSES) schools, $t(9) = 6.425$, $P < 0.05$, and $t(19) = 3.969$, $P<0.05$ respectively.

Tree Tops (HSES) also scored significantly higher than Mumuni and Twatasha; $t(19) = 4.450$, $P< 0.05$, and $t(19) = 6.700$, $P <0.05$ respectively.

Whereas there was no significant difference in performance between Lake Road and Tree Tops, there was a significant difference between Twatasha and Mumuni; $t(19) = 2.328$, $P<0.05$.

When the pretest and postest scores for both the treatment and control groups were compared to test the hypothesis that performance on the special paper test improves with direct teaching, the following results, as shown in table 4, were produced.

Table 4: Pretest and posttest scores for the treatment and control groups compared.

SCHOOL		PRETEST			POSTTEST			t VALUE	CRITICAL VALUE = 0.5
		NO. of Subjects	SP mean score %	SP score S.D.	NO. subjects	SP score %	SP S.D.		
TREATMENT GROUP	Tree Tops	20	71	7.18	20	77	9.82	2.206	2.093
	Twatasha	20	40	20.20	15	43	22.88	0.403	2.145
CONTROL GROUP	Mumuni	20	54	17.76	18	61	23.15	1.227	2.110
	Lake Road	20	75	11.40	18	80	13.65	1.218	2.110

Tree Tops, a treatment group school, showed a significant improvement in the posttest; $t(19)=2.206$ $P < 0.05$.

Twatasha also a treatment group school, however, did not show a significant improvement; $t(14)=0.403$, $P<0.05$.

Mumuni and Lake Road, which composed the control group both did not show statistically significant improvement at 0.05 level.

The scrutiny of the 1986 special paper to determine the items' suitability to measure scholastic aptitude yielded a description of the test items and their categorisation according to Bloom's taxonomy of learning outcomes as is shown in tables 5a and 5b below.

Table 5a: Description of the special paper test items.

PAPER ONE		PAPER TWO	
TYPE OF ITEM	NUMBER OF ITEMS	TYPE OF ITEM	NUMBER OF ITEMS
Opposites	3	Pattern completion	5
Word pattern	1	Odd one out	6
Letter series	2	Analogy	33
Number series	2	Progression	6
Analogy	14		
Odd one out	9		
Always has	8		
Alphabetical order	5		
Problems	6		
TOTAL	50	TOTAL	50

Paper one comprised 9 classes of item type. These were: Opposites, word pattern, letter series, always has, analogy, odd one out, number series, alphabetical order, and problems. Paper two comprised four classes that include pattern completion, odd one out, analogy, and progression. The range of frequency for each item type was from 3 to 9, and 5 to 33 for paper one and paper two respectively.

Table 5b: Categorisation of special paper test item types according to Bloom's taxonomy of learning outcomes.

Description of Major Cognitive Domains of Bloom's Taxonomy of Learning Outcomes.	Illustrative verbs for stating specific learning	Corresponding Type of Test item.
1. **Knowledge.** The remembering of previously learned material ranging from specific facts to complete theories. Requires only recall of appropriate information. Lowest level of learning outcomes.	Defines, defends identifies labels, lists, matches, names outlines. reproduces, selects, states.	-opposites -letter series -alphabetical order
2. **Comprehension.** The ability to grasp the meaning of material shown by translating material from one form to another, by interpreting material, and by estimating future trends. Learning outcomes go a step beyond the simple remembering of material, and represent the lowest level of understanding	Converts, defends distinguishes, estimates, explains, extends generalises, gives examples infers, paraphrases, predicts, rewrites, summarises.	-word pattern -odd one out -pattern completion -progression -analogy
3. **Application.** The ability to use learned material in new and concrete situations. May include the application of such things as rules, methods, concepts,principles laws, and theories. Learning outcomes in this area require a higher level of understanding than those under comprehension.	Changes, computes demonstrates, discovers manipulates, modifies, operates, predicts, relates, shows solves, uses.	-number series
4. **Analysis.** The ability to break down material into its component parts so that its organisational structure may be understood. This may include the identification of the parts, analysis of the relationships between	Breaks down, diagrams, differentiates discriminates, distinguishes, identifies. illustrates, infers,	-problems -always has

parts, and recognition of the organisational principles involved. Learning outcomes represent a higher intellectual level than comprehension and application because they require an understanding of both the content and the structural form of the material.	outlines points out, relates, selects separates, subdivides.	
5. **Synthesis.** Ability to put parts together to form a whole. May involve the production of a unique communication, a plan of operations, or a set of abstract relations. Learning outcomes stress creative behaviour-formulation of new patterns or structures.	Categories, combines, compiles, composes, creates, devises, designs, explains, generates, modifies, organises, plans, arranges, reconstructs, relates, recognises, revises, rewrites, summarises tells writes	NONE
6. **Evaluation.** Ability to judge the value of material for a given purpose-internal or external criteria to be determined or provided. Learning outcomes are highest in the cognitive hierachy because they contain all elements of the other categories.	Appraises, compares, concludes, critises, describes, discriminates, explains, justifies, interprets, relates, summarises, supports.	NONE

The analysis of the 1986 special paper scores for the top 10 pupils from the high ability group (U), and the bottom 10 from the low ability group (L), produced results as shown in tables 6a and

6b. This was necessary to determine the test items' difficulty level and discrimination.

Table 6a: Item difficulty and discrimination power for the Special Paper One test.

			ALTERNATIVES					INDEXES	
ITEM	PUPILS		A	B	C	D	OMIT	DF%	DIS 1
1	U	10	1	0	9	0	0	65	0.50
	L	10	2	1	4	0	0		
2	U	10	0	0	1	9	0	70	0.60
	L	10	0	2	4	4	0		
3	U	10	0	0	1	9	0	75	0.30
	L	10	1	2	1	6	0		
4	U	10	0	0	0	0	0	74	0.63
	L	10	2	4	0	3	1		
5	U	10	10	0	0	0	0	85	0.30
	L	10	7	0	2	1	0	(73.8)	
6	U	10	0	10	0	0	0	75	0.53
	L	10	1	5	2	1	1		
7	U	10	0	0	10	0	0	58	0.95
	L	10	3	3	1	2	1		
8	U	10	1	8	0	1	0	60	0.42
	L	10	1	4	1	1	0		
9	U	10	0	0	10	0	0	73	0.63
	L	10	3	1	4	1	1		
10	U	10	0	0	9	1	0	63	0.63
	L	10	0	2	4	0	0		
11	U	10	9	1	0	0	0	57	0.74
	L	10	2	1	2	4	1		
12	U	10	0	0	0	10	1	68	0.74
	L	10	1	2	3	3	1		

		ALTERNATIVES					INDEXES	
ITEM	PUPILS	A	B	C	D	OMIT	DF%	DIS 1
13	U 10	9	0	1	0	0	52	0.84
	L 10	1	4	2	2	1		
14	U 10	10	0	0	0	0	68	0.74
	L 10	3	3	2	1	1		
15	U 10	0	0	0	10	0	57	0.95
	L 10	0	3	5	1	1		
16	U 10	0	0	10	0	0	68	0.74
	L 10	3	1	3	2	1		
17	U 10	10	0	0	0	0	68	0.74
	L 10	3	3	2	1	1		
18	U 10	0	0	0	10	0	65	0.70
	L 10	1	4	2	3	0		
19	U 10	0	6	3	1	0	42	0.42
	L 10	4	2	1	2	1		
20	U 10	10	0	0	0	0	57	0.95
	L 10	1	2	3	3	1		
21	U 10	1	0	0	9	0	52	0.84
	L 4	2	2	4	1	1		
22	U 10	0	1	7	2	0	63	0.21
	L 10	3	0	5	1	1		
23	U 10	0	7	3	0	0	63	0.22
	L 10	2	5	1	0	2		
24	U 10	0	0	10	0	0	68	0.74
	L 10	2	1	3	3	1		
25	U 10	0	10	0	0	0	78	0.53
	L 10	3	5	0	1	1		
26	U 10	1	0	0	9	0	72	0.56
	L 10	0	1	3	4	0		
27	U 10	1	1	8	0	0	47	0.74
	L 10	2	2	1	2	1		
28	U 10	7	0	1	2	0	63	0.74
	L 10	5	1	3	0	1		

ITEM	PUPILS		ALTERNATIVES A	B	C	D	OMIT	INDEXES DF%	DIS 1
29	U	10	0	10	0	0	0	63	0.84
	L	10	4	2	2	1	1		
30	U	10	0	0	10	0	0	63	0.84
	L	10	2	1	2	4	1		
31	U	10	0	0	0	10	0	63	0.84
	L	10	2	2	3	2	1		
32	U	10	0	0	9	1	0	58	0.74
	L	10	2	3	2	2	1		
33	U	10	7	1	0	1	1	50	0.50
	L	10	2	1	4	2	1		
34	U	10	0	0	9	1	0	55	0.78
	L	10	2	4	2	2	0		
35	U	10	0	10	0	0	0	58	0.95
	L	10	4	1	4	0	1		
36	U	10	0	0	0	10	0	68	0.74
	L	10	1	3	2	3	1		
37	U	10	0	10	0	0	0	68	0.74
	L	10	1	3	2	3	1		
38	U	10	10	0	0	0	0	53	1.00
	L	10	0	4	1	4	1		
39	U	10	0	0	4	6	0	32	0.63
	L	10	1	7	1	0	1		
40	U	10	2	0	8	0	0	42	0.84
	L	10	1	6	0	2	1		
41	U	10	10	0	0	0	0	68	0.74
	L	10	3	2	2	2	1		
42	U	10	0	1	8	1	0	53	0.63
	L	10	4	1	2	2	1		
43	U	10	1	0	0	9	0	63	0.63
	L	10	0	4	2	3	1		

ITEM	PUPILS		ALTERNATIVES A	B	C	D	OMIT	INDEXES DF%	DIS 1
44	U	10	1	1	2	6	0	36	0.53
	L	10	1	0	7	1	1		
45	U	10	0	1	2	7	0	47	0.53
	L	10	3	3	1	2	1		
46	U	10	0	0	10	0	0	68	0.74
	L	10	3	2	3	1	1		
47	U	10	1	9	0	0	0	53	0.84
	L	10	4	1	2	2	1		
48	U	10	0	4	5	1	0	32	0.22
	L	10	2	2	1	4	1		
49	U	10	0	0	1	9	0	68	0.53
	L	10	2	0	3	4	1		
50	U	10	10	0	0	0	0	74	0.63
	L	10	4	2	0	3	1		
						MEAN =		61	0.66

Table 6b: Item difficulty and item discrimination power for the 1986 Special Paper Two test.

ITEM	PUPILS		ALTERNATIVES A	B	C	D	OMIT	INDEXES DF%	DIS 1
1	U	10	0	0	0	10	0	70	0.6
	L	10	3	1	2	4	0		
2	U	10	0	0	10	0	0	55	0.9
	L	10	2	1	1	6	0		
3	U	10	1	0	4	5	0	25	0.5
	L	10	4	3	3	0	0		
4	U	10	9	0	1	0	0	50	0.8
	L	10	1	3	2	4	0		
5	U	10	0	0	9	0	1	63	0.6
	L	10	4	2	3	1	0		
6	U	10	8	1	0	1	0	60	0.4
	L	10	4	3	3	0	0		
7	U	10	0	0	0	10	0	70	0.6
	L	10	1	3	2	4	0		
8	U	10	10	0	0	0	0	55	0.9
	L	10	1	3	4	2	0		
9	U	10	0	10	0	0	0	70	0.74
	L	10	1	4	1	4	0		
10	U	10	0	0	10	0	0	75	0.50
	L	10	1	1	5	3	0		
11	U	10	0	0	0	10	0	60	0.8
	L	10	4	2	2	2	0		
12	U	10	0	0	10	0	0	75	0.5
	L	10	3	0	5	2	0		
13	U	10	0	10	0	0	0	55	0.9
	L	10	3	0	4	2	0		
14	U	10	1	7	1	1	0	55	0.3
	L	10	1	4	1	4	0		
15	U	10	8	2	0	0	0	40	0.8
	L	10	0	2	4	4	0		

ITEM	PUPILS		ALTERNATIVES A	B	C	D	OMIT	INDEXES DF%	DIS 1
16	U	10	0	0	0	10	0	55	0.9
	L	10	1	6	2	1	0		
17	U	10	9	0	0	1	0	65	0.5
	L	10	4	4	0	2	0		
18	U	10	0	0	0	10	0	65	0.7
	L	10	2	1	4	3	0		
19	U	10	0	10	0	0	0	80	0.3
	L	10	2	7	1	0	0		
20	U	10	2	8	0	0	0	55	0.5
	L	10	2	3	3	2	0		
21	U	10	0	0	0	10	0	50	1.0
	L	10	3	2	5	0	0		
22	U	10	0	6	3	1	0	50	0.4
	L	10	4	2	4	0	0		
23	U	10	1	9	0	0	0	60	0.6
	L	10	4	3	1	2	0		
24	U	10	0	1	9	0	0	60	0.6
	L	10	3	2	3	2	0		
25	U	10	0	0	10	0	0	50	0.8
	L	10	3	3	2	2	0		
26	U	10	10	0	0	0	0	65	0.7
	L	10	3	5	1	1	0		
27	U	10	0	10	0	0	0	75	0.5
	L	10	3	5	1	1	0		
28	U	10	9	1	0	0	0	60	0.6
	L	10	3	1	1	5	0		
29	U	10	8	2	0	0	0	55	0.5
	L	10	3	2	2	3	0		
30	U	10	0	3	7	0	0	55	0.3
	L	10	2	4	4	0	0		

ITEM	PUPILS		A	B	C	D	OMIT	INDEXES DF%	INDEXES DIS 1
31	U	10	2	0	0	8	0	50	0.6
	L	10	2	3	3	2	0		
32	U	10	1	8	0	1	0	55	0.5
	L	10	0	3	5	2	0		
33	U	10	0	7	3	0	0	40	0.6
	L	10	3	1	4	2	0		
34	U	10	8	2	0	0	0	50	0.6
	L	10	2	4	1	3	0		
35	U	10	1	0	1	8	0	40	0.8
	L	10	2	4	4	0	0		
36	U	10	0	10	0	0	0	65	0.7
	L	10	2	3	3	2	0		
37	U	10	7	1	1	1	0	50	0.4
	L	10	3	2	1	4	0		
38	U	10	0	0	10	10	0	80	0.4
	L	10	3	0	6	3	1		
39	U	10	0	0	9	1	0	65	0.5
	L	10	2	1	4	3	0		
40	U	10	0	0	0	10	0	75	1.5
	L	10	1	2	2	5	0		
41	U	10	1	2	5	2	0	35	0.3
	L	10	2	3	2	3	0		
42	U	10	2	0	1	7	0	45	0.5
	L	10	5	1	2	2	0		
43	U	10	0	3	6	1	0	35	0.5
	L	10	3	3	1	3	0		
44	U	10	0	2	7	0	1	53	0.4
	L	10	4	0	3	3	0		
45	U	10	1	6	0	3	0	35	0.5
	L	10	1	1	4	4	0		

ITEM	PUPILS	ALTERNATIVES A	B	C	D	OMIT	INDEXES DF%	DIS 1
46	U 10	1	1	3	5	0	45	0.1
	L 10	4	1	1	4	0		
47	U 10	0	9	1	0	0	55	0.7
	L 10	4	2	3	1	0		
48	U 10	1	0	9	0	0	45	0.9
	L 10	6	3	0	1	0		
49	U 10	0	0	5	5	0	40	0.2
	L 10	2	1	4	3	0		
50	U 10	7	0	2	1	0	40	0.6
	L 10	1	3	3	3	0		
				MEAN =			50	0.6

Item difficulty for Paper One was 61 per cent on average, with items ranging from 32 per cent to 85 per cent. Average discrimination power was 0.7, with items ranging from 0.3 to 1.0. Paper Two had a difficulty mean of 56 per cent, with items ranging from 0.1 to 1.0.

A further analysis of item difficulty enabled the placement of items in respect of the level of difficulty. This was done to determine whether or not the paper meets the requirement of placing easy items first and difficulty ones last. These results are shown on figure 1 and figure 2 below.

Figure 1: Special Paper One test items analysed according to difficulty level

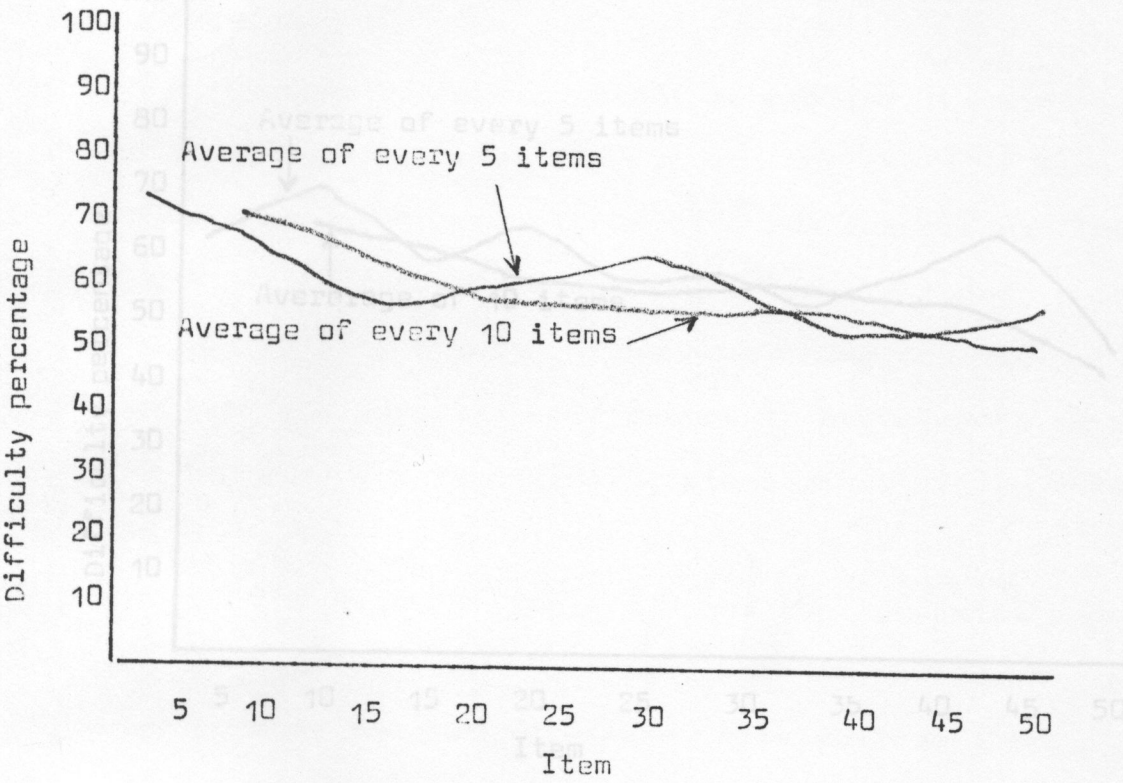
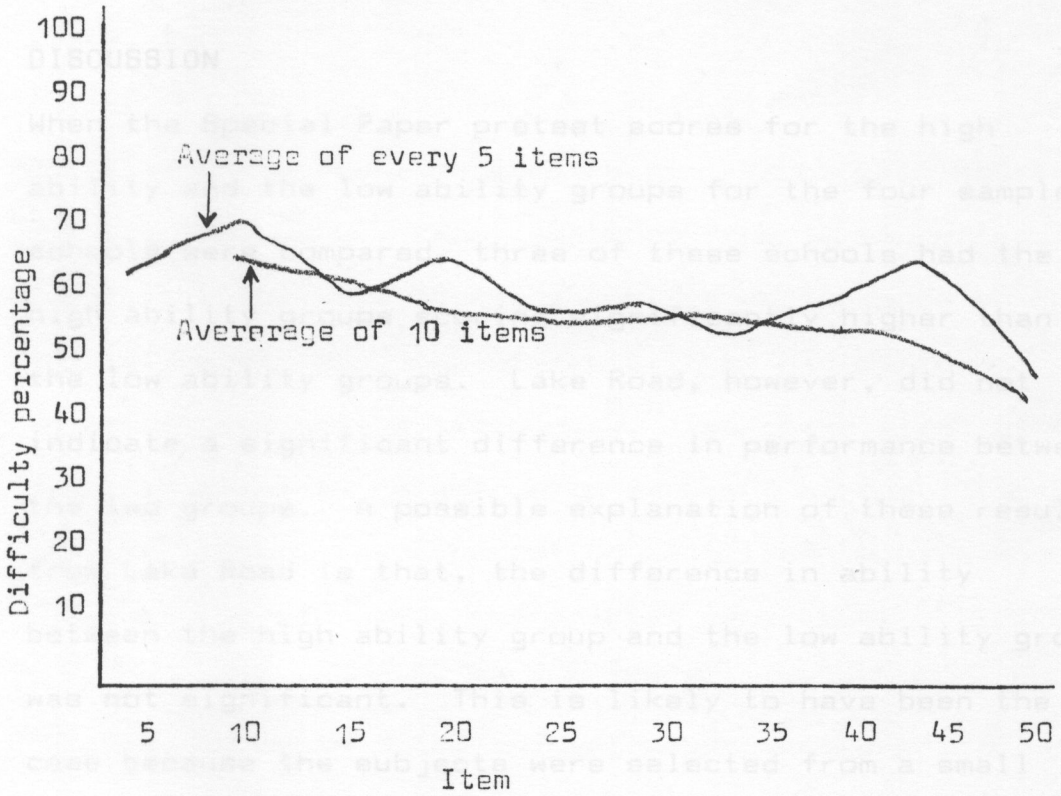


Figure 2: Special Paper Two test items analysed according to difficulty level.



According to the formular used (dif. = $\frac{R \times 100}{T}$), the higher the difficulty percentage, the easier the item.

CHAPTER FIVE

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

DISCUSSION

When the Special Paper pretest scores for the high ability and the low ability groups for the four sample schools were compared, three of these schools had the high ability groups scoring significantly higher than the low ability groups. Lake Road, however, did not indicate a significant difference in performance between the two groups. A possible explanation of these results from Lake Road is that, the difference in ability between the high ability group and the low ability group was not significant. This is likely to have been the case because the subjects were selected from a small class of 24 pupils with a narrow range in performance. The other schools, on the other hand, selected their subjects from much larger pupil populations. For instance, Twatasha, Mumuni, and Tree Tops selected their subjects from populations of about 300,160 and 80 pupils respectively.

These results supported the hypothesis that the higher a pupil's performance in primary school subjects, the higher will be his score on the special paper test. They seem to suggest that high ability pupils will perform better on the special paper test than the low ability ones. Theoretically, these findings may imply

that there is a positive correlation between performance in school subjects and performance on the special paper test.

Correlation between the 1985 Special Paper test scores and the 1987 JSSLE, from Lake Road School, supported the hypothesis that the higher a pupil's performance on the special paper test, the higher will be his performance in school subjects at secondary. The correlation of 0.8 demonstrated this relationship as being strong. These results support the expectation that those pupils who produce high scores on the special paper test will tend to produce high scores on the JSSLE at secondary school. This may imply that the special paper test predicts scholastic achievement.

The difficulty level of 58 percent and the discrimination power of 0.68 apparently indicate that the paper is psychometrically functioning well. This is so because a suitable test for assessing a given attribute should produce a difficulty level of 50 percent and a discrimination power of 0.5 percent. Also, the succession of difficulty in the placing of test items indicated a general trend of easy items being first and difficult ones last, which is a requirement for an aptitude test.

However, a further scrutiny of test items in respect of subject matter and involvement of required mental functions, and the construction of the items displayed a less supportive picture.

In regard to subject matter and involvement of required mental functions, test items were assessed according to Bloom's taxonomy of learning outcomes, and the results indicated that the items included only the knowledge, comprehension, application, and analysis levels, leaving out the synthesis and the evaluation levels. This weakened the test because the demonstration of the ability to plan operations or to reason about sets or abstract relationships which is involved in the synthesis stage (level), and the ability to judge the value of material for a given purpose which is involved in the evaluation stage are essential aspects of scholastic aptitude. Without these two, the test only assesses low level learning outcomes. It should be noted also that most items were at the levels of knowledge and comprehension (78 out of 100 items), a condition that makes the test less representative of the tasks in which scholastic aptitude is appropriately demonstrated.

Apart from assessing only a low level of learning outcomes, the nature of items in some cases required the

pupils to have particular skills or particular knowledge, the acquisition of which may depend on exposure to a particular culture or environment. Still, in some cases, the construction of the items was erroneous such that an item had no correct answer from the multiple choice given, or there was more than one correct answer. We shall discuss each one of these cases in respect of subject matter and construction (clarity) taking case examples.

Knowledge

Opposites: Adore is to Hate as Come is to _____

A Came

B Go

C Went

D Disappear

This type of item, as concerns subject matter, requires understanding of the meaning of the words and the relationship between the alternative answers. This understanding, however, may only involve matching of the words, identifying or remembering of previously learned material. The problem of unfamiliar vocabulary may arise here thereby inhibiting performance. Performance on such an item may thus be influenced more by lack of understanding the question than lack of aptitude. If such items are included in the test, they should be

checked for appropriate vocabulary otherwise they will reduce the effectiveness of the test. On average, opposites had an item difficulty of 65 percent and a discrimination power of 0.6.

Letter series: Which letter in the word "christmas" would come second if all the words were placed in alphabetical order?

A a

B i

C c

D h

Alphabetical order: If the following words were put in alphabetical order, which one would come second?

A Puzzle

B Steel

C Punch

D Punt

The letter series and the alphabetical order types of item both require remembering of the alphabet and the alphabetical placing of words. This type of item does not require the pupil to demonstrate any ability other than remembering specific facts or formulae. If the pupil has no prior knowledge of the principles guarding

the putting of words in alphabetical order, he may fail to answer the item correctly. This may not necessarily demonstrate lack of aptitude. Such items therefore, are less suitable for assessing scholastic aptitude. If anything, they may be useful in assessing scholastic achievement. These types had on average a difficulty level of 42 percent and a discrimination power of 0.7. This implies that pupils tend to find items of this type difficult.

Comprehension

Word pattern: Which word will complete the pattern?

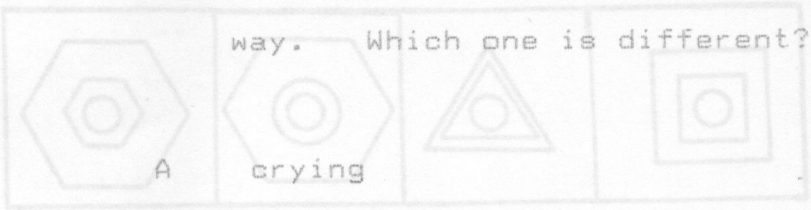
NONE	ONE
FOLD	OLD
BROAD	ROAD
REAR	_____

A	EARN
B	READ
C	EAR
D	EARTH

This type of item requires the ability to deduce information from available trends. To tackle the item successfully, the pupil should grasp the meaning of the material which is shown. Vocabulary problems do not

arise here. Also, the item does not require prior knowledge of specific facts. Therefore, the item is appropriate for assessing aptitude, and as such, appropriate for the special paper to assess scholastic aptitude. This type of item had on average a difficulty level of 65 percent and a discrimination power of 0.7.

Odd one out: Three of the following are alike in some way. Which one is different?



A crying

B wailing

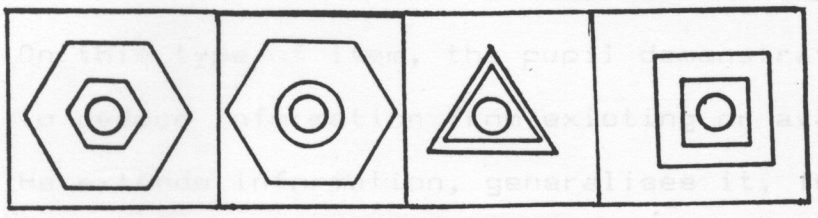
C moaning

D laughing

To answer this type of item, the pupil should be able to understand the meaning of the words and the relationships that they have. By so doing, the pupil demonstrates his ability to grasp, compare and contrast the given information. To do this, the pupil does not need prior knowledge of specific theories of learned material. However, this type of item may raise the problem of vocabulary. For the pupil to be able to compare and contrast the information, he has to understand what the words mean. As alluded earlier,

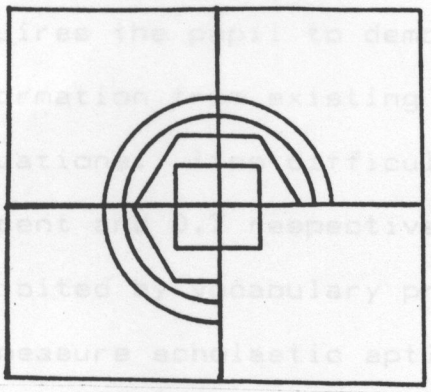
when unfamiliar vocabulary is used in such items, the test tends to assess one's vocabulary and not his scholastic aptitude.

Odd one out: Which one is different?



This type of item, like the verbal one, requires the pupil to demonstrate his ability to compare and contrast information. Because non-verbal items do not raise vocabulary problems they may be appropriate for assessing scholastic aptitude. In the present study, items of this type had, on average, the level of difficulty of 55 percent and the discrimination power of 0.4. This shows that they appeared to be appropriate in assessing what they were supposed to assess.

Pattern Completion:



A

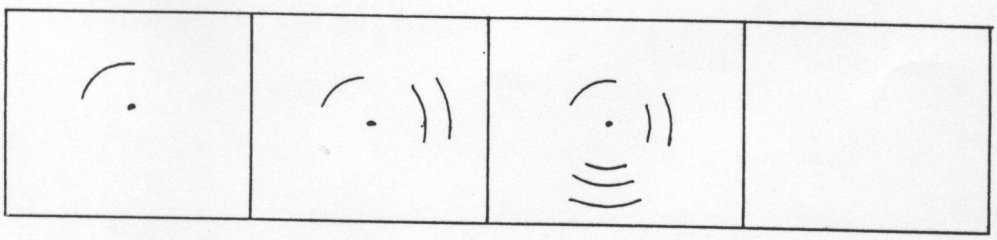
B

C

D

On this type of item, the pupil demonstrates his ability to deduce information from existing or available trends. He extends information, generalises it, infers from it and predicts future situations. The item does not require the pupil to have specific prior facts or theories. Since the item is non-verbal, the pupil's required behaviour is not inhibited by a break down in communication. Such items are therefore appropriate for assessing scholastic aptitude. On average, this type of items had a difficulty level of 50 percent and a discrimination power of 0.9.

Progression:



A

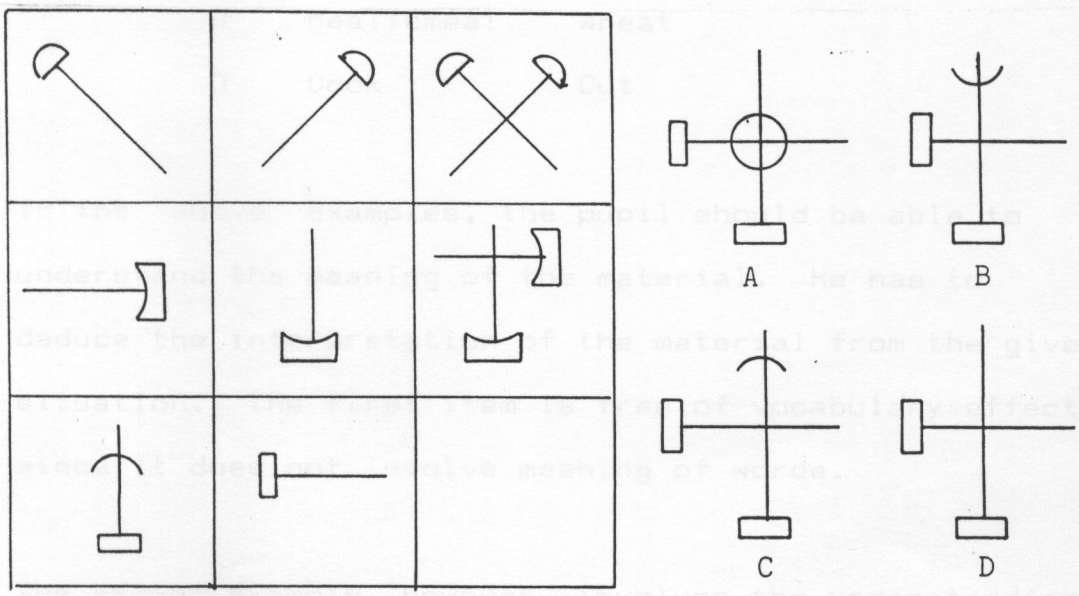
B

C

D

Like the pattern completion items, the progression item requires the pupil to demonstrate his ability to deduce information from existing trends and predict future situations. Item difficulty and discrimination was 60 percent and 0.7 respectively. Performance on it is not inhibited by vocabulary problems, and hence it is likely to measure scholastic aptitude.

Analogy:



This type of item is like the progression and pattern completion types. It requires deducing of information from existing trends. It is also non-verbal and hence

free of vocabulary effects. It had an item difficulty of 60 percent and a discrimination power of 0.6.

Analogy: If ENGLISH is 9820345 then HEN is .....

A 390

B 253

C 895

D 598

..... is to Nshima as Flour is to.....

A	Tomato	Tea
B	Mealiemeal	Bread
C	Mealiemeal	Wheat
D	Cook	Cut

In the above examples, the pupil should be able to understand the meaning of the material. He has to deduce the interpretation of the material from the given situation. The first item is free of vocabulary effects since it does not involve meaning of words.

The second example, however, involves the understanding of meanings of words thereby making performance on it affected by vocabulary. It also involves cultural setting. Nshima is a food made from ground maize (mealie-meal, or ground cassava roots (cassava meal) commonly used by native Zambians. Bread is also a food, made from ground wheat (flour). To rural Zambian

children, it may be difficult to know that bread is made from flour, let alone knowing what bread is. Also, to urban children, especially immigrant groups, it may be difficult to know what mealie-meal or nshima is. This item is therefore not suitable for an aptitude test.

Application

Number series: Which number comes next?

81 - 64 - 49 - ....

A 36

B 33

C 25

D 27

To be able to answer this type of item, the pupil should be able to interpret the material involved. He should discover the relationship between the figures and then apply the rules governing this relationship. The number series items had difficulty level of 50 percent and a discrimination power of 0.5. For a scholastic aptitude test, this was ideal. The effect of such good items on the quality of the paper was however small because there were only two of them.

Analysis

In answering this type of items, the pupil demonstratee his ability to understand and analyse the relationship between the parts, and the ability to recognise the organisational principles involved. Since this type of verbal communication does not include names of objects, the vocabulary problem is minimal, and as such these items are appropriate for scholastic aptitude tests. Moreover, the level of learning outcomes involved is reasonably high.

Always has: A tree always has .....

A leaves

B flowers

C roots

D fruits

On this type of items the pupil should understand the qualities of the subject concept as well as the relationship between the concept and each of the comparison concepts. To do this, he must have the ability to analyse. So long as the concepts used do not pose cultural or environmental problems, this type of items is ideal for a scholastic aptitude test. These items had a difficulty level of 62 percent and a discrimination power of 0.6, which is reasonably close to the ideal.

The scrutiny of test items also indicated that in some cases items were either ambiguous, had no correct answer or they were biased. The following are the cases as cited from past examination question booklets.

No correct answer given:

Problems: Mwango is ten years older than his wife. His wife was 18 years old when their first child was born. The child is now 19 years old. How old is Mwango now?

A 37

B 45

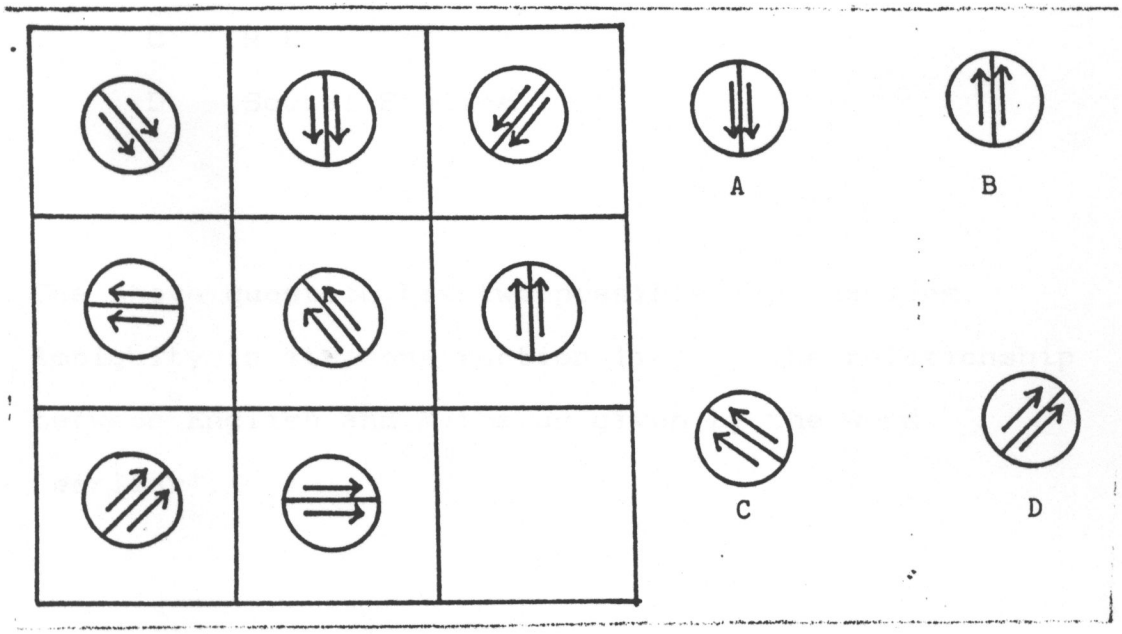
C 47

D 50

In answering this type of items, the pupil demonstrates his ability to understand and analyse the relationship between the parts, and the ability to recognise the organisational principles involved. Since this type of verbal communication does not include names of objects, the vocabulary problem is minimal, and as such these items are appropriate for scholastic aptitude tests. Moreover, the level of learning outcomes involved is reasonably high.

The scrutiny of test items also indicated that in some cases items were either ambiguous, had no correct answer or they were biased. The following are the cases as cited from past examination question booklets.

No correct answer given:



On this item, there is no correct answer given. In this case, the pupil is left with no alternative but to guess. Guessing, obviously, is not what the test is meant for.

The answer is supposed to be the figure in the top left hand box.

More than one correct answer possible:

The time table showed that the Grade 7s would have English earlier than Maths. Social Studies would come next followed by R.E. which would come later than Science. Which subject would come second from the last?

A English
B Science
C R.E.
D Social Studies

The above question has two possible explanations. Ambiguity in its construction lies in the relationship between English and Maths as given by the word 'earlier'.

The relationship implied is the same as for the other subjects as stated in the following example:

English came earlier than any other subject.

In this example, Maths can either be placed next to English, or at the end of the series. If we choose to place it next to English, the succession will be English , Maths, Social Studies, Science, and R.E., making B as the correct answer. However, if we place it at the end of the series, the succession becomes English, Social Studies, Science, R.E., and Maths and C becomes the correct answer.

More than one correct answer possible:

Three of the following are alike in some way. Which one is different?

A door

B bed

C chair

D car

On this item, the pupil can have either 'car' as the exception or he can have 'door' depending on what he bases his classification. He can on the one hand, have

'door' if he bases his classification on the fact that a door may be not a whole but a part. For instance, a door can be part of a house or a car whereas the rest of the objects are wholes. On the other hand, he can argue that 'car' is different because it is mobile and it is not for indoor use.

Such ambiguity in test items should be avoided in a multiple choice type of test. It is unfair to fail a pupil only because he has used a different interpretation from the examiner's when he has convincing reasons for his answer. Ambiguity as expressed here may make the special paper a less valid instrument for assessing scholastic aptitude.

Socio-economic status biased item:

RADIO is to TUNE as TELEPHONE is to.....

A RING

B CALL

C DIAL

D TALK

Items which involve objects which are not in common use among members of some socio-economic groups put some pupils at a disadvantage. For example, in Zambia, rural

inhabitants do not have telephones in their homes. As a result, rural pupils would find items such as the one given above, difficult.

Whereas item difficulty and item discrimination results may have indicated that the special paper test was appropriate for assessing scholastic aptitude, the comparison of results from the treatment and the control groups, and the analysis of performance according to socio-economic status groups introduced doubt. The comparison of results between the treatment and the control groups indicated that the paper is sensitive to coaching. Sensitivity to coaching makes the paper less valid for assessing scholastic aptitude.

Administratively, although test items are not supposed to be exposed to the pupils except at examination times, some schools can privately teach items to pupils whereas some schools may not. In this case, those pupils who have the chance to practice taking the test items will have better chances of doing well on the test than those pupils who do not. In fact, it is not uncommon to find special paper copies in schools and teachers coaching their pupils. Unless it is administratively ensured that past examination copies do not get into unauthorised circulation, the test will not be a fair measure of scholastic aptitude.

The analysis of pupil performance in respect of socio-economic status indicated that the test discriminates against socio-economic groups. Pupils from the high socio-economic status performed better than the pupils from the low socio-economic status group. This implies that the test may not be measuring scholastic aptitude but some other trait that is influenced by environmental conditions. The use of the special paper for selection purposes may therefore promote inequality in the provision of education. Specifically, the low socio-economic status group is put at a disadvantage when competing for secondary school places.

CONCLUSION

From the findings of the present study, the following conclusions can be made:

1. Although performance on the special paper test correlates with performance in school subjects both at primary and secondary school levels, the paper may not assess scholastic aptitude. This is so because the paper showed that it is influenced by coaching and socio-economic background, factors to which a scholastic aptitude test should not be susceptible as they provide evidence that the test measures some other traits.

2. The special paper test can not be an appropriate instrument for measuring scholastic aptitude. This is so because the test items are not free of ambiguity and the subject matter is of low level, making the test less representative of the mental functions that are indicative of scholastic aptitude.

RECOMMENDATIONS

On the basis of the present study's findings, the following recommendations are made:

1. The special paper should be taught in school like any other subject. This recommendation is made on the basis that the paper is susceptible to coaching. Since it is difficult, if not impossible to ensure that no pupil can have the change to practice on the test items prior to the examination, if the paper is allowed to be taught in school, then fairness will be ensured to all pupils.

   As regards the tests' susceptibility to socio-economic background, it is recommended that the Ministry of Education (the Examinations Council) continues varying the cut-off points according to regions. In this case, special attention should be paid as to what the socio-economic status of a particular school is. That is, the higher the socio-economic status of the school, the higher should be the cut-off-point mark.

2. The quality of the special paper should be improved. Whatever traits the special paper may measure, it can not be efficient if the subject matter and the construction of test items remains at the present standard. The subject matter should be improved such that it includes the synthesis and evaluation levels of learning outcomes. This may be achieved by involving in the construction of test items, other specialised branches of Psychology personnel other than the present practice of the Examinations Council personnel.

   The present shortcomings in the construction of test items is as a result of the reproduction of old items. The involvement of a wider range of specialised personnel would easily rid the paper of culture bias. To rid the test of ambiguity it only requires a thorough proof reading (editing) of the test items.

3. The special paper should cease to be treated as a scholastic aptitude test. Until the recommendations made above are implemented, and validity retested, the test should not be treated as a scholastic aptitude test. When, however, the recommendations in (2) above have been met, the

test may be used as a compulsory selection subject like English. In the mean time, the paper may be treated as an optional selection subject.

Below is an example of a test item that would include the synthesis and the evaluation levels of learning outcomes. The material is an extract from T.G.R. Bower's (1979) work on adolescent and adult thought.

Instructions

Read the following statements carefully and answer the questions that follow.

1. A right action is an action that will bring about at least as much good, or, failing to do that, will avoid at least as much evil as any other action open to the agent at the time of acting.

2. A good man is a man who always does what seems to him, after due consideration, to be right.

3. It is always wrong to tell a lie or break a promise.

4. Suffering itself is undoubtedly evil.

5. In some cases it seems obvious that the only consequence of telling the truth or keeping a promise will lead to more suffering than would result from the opposite behaviour.

Questions.

1. Are these statements compatible, one with another?

2. If not, what is the least number that must be rejected to yield a completely consistent set?

3. Write out such a list, containing the fewest possible rejections, and state briefly wherein lies the incompatibility between those you reject and those you retain.

As alluded to earlier, an attempt to test scholastic aptitude through the use of multiple choice type of test may not be appropriate since some required demonstrations are expressive. Scoring on an item that includes expressive behaviour would not be 'right or wrong' but 'how much of the behaviour indicative of the required trait is exhibited'.

BIBLIOGRAPHY

1. BALDWIN Alfred, L., (1967) Theoreis of Child Development. John Wiley and Sons Inc. U.S.A. Card No. 66 - 26733.

2. BARBE Walter, B. (1965) Psychology and Education of the Gifted: Selected Readings. Appleton - Century - Crofts, Meridith Publishing Company, New York. Card No. 65 - 12346.

3. BERNARD Harold, W. (1964), Psycholoy of Learning and Teaching. MacGraw-Hill, Inc. New York, St. Louis, San Franciasco, Toronto, London, Sydney. Catalogue No. 64 - 8726.

4. BORG W.R. and ALL M.D. (1979), Educational Research, Third Edition, Longman Inc. New York

5. BOWER, Gordon, H. and HILGARD eRNEST, r. (1981). Theories of Learning: Fifth Edition. Prentice-Hall Inc. Englewood Cliffs, N.J. 07632. ISBN 0 - 13 - 914432.

6. BOWER, T.G.R. (1979), Human Development, W.H. Freeman and Company, San Francisco, U.S.A. ISBN 0 - 7167 - 0058 - 1.

7. CATELL Raymond, B. and BUTCHER John, H. (1968), The Prediction of Achievement and Creativity. The Bobbs - Merill Company, Inc. Indianapolis, New York. Catalo No. 67-18662.

8. CLARIZIO Harvey, F. CRAI Robert, and MEHRENS Willians, A. (1981), Contemporary Issues in Educational Psychology: Fourth Edition. Allyn and Balcon, Inc. 470 Atlantic Avenue, Beston. ISBN 0-205 - 07331 - X.

9. COLADARACI, Arthur, P. (1955), Education Psycholoy. Dryden Press Professional Books in Education. The Dryden Press, Inc. New York 19.

10. CRONBACK Lee, J. (1984), Essentials of Psychological Testing: Fourth Edition. Haper and Row Publishers Inc. Haper International Edition. ISBN 0 - 6 - 350249 - 6.

11. DIGGORY Sylvia Farnham, (1972), Cognitive Processes in Education: A Psychological Preparation for Teaching and Curriculum Development. Harper and Row, Publishers, New York, Evanston, San Francisco, London. Catalog No. 7- - 178105.

12. DYER, J.R.. (1979), Understanding and Evaluating Educational Research. Addison-Wesley Publishing Company, Reachin, Massachusetts, Menlo Park, California, London, Amsterdam, Don Mills, Ontario, Sidney, ISBN 0 -201 - 01184 - 0.

13. FLAVELL John, H. (1977), Cognitive Development. Prentice-Hall, Inc. Englewood Cliffs, New Jersey. ISBN 0 - 13 - 139766 - 4 pbk.

14. GALL Meredith, D. and WARD Beatrice A. (1974), Critical Issues in Educational Psycholoy. Little, Brown and Company Inc. Boston Toronto. Catalog. No. 73 - 20936.

15. GRONLUND Norman, E (1981), Measurement and Evaluation in Teaching: 4th Edition, MacMillan Publishing Company Inc. New York. ISBN 0 - 02 - 348020 - 3.

16. INGENKAMP Kartheinz, (1969), Developments in Educational Testing: Volume 2. University of London Press Ltd., St. Paul's House, Warwick Lane, London, EC4. ISBN 340 - 9637 - 3.

17. MACGINITILE Walter, H. and BALL Samuel, (1968), Readings in Psychological Foundations of Education. McGraw-Hill, Inc. New York, St. Louis, San Francisco, Toronto, London, Sydney. Catalog No. 68 - 14339.

18 MACMILLAN J.H. and Schumacher, S. (1984), Research in Education. Little, Brown and Co. Boston, Toronto.

19. NOLL V.H. and Scannell D.P. (1972), Introduction to Educational Measurement. Third Edition. Houston Mifflin Co. Boston.

20. NUNNALLY Jum, C. (1978), Psychometric Theory: Second Edition. McGraw-Hill Book Company, New York, U.S.A. ISBN 0 - 07 - 047465 - 6.

21. PIAGET J. (1947), The Pscyhology of Intelligence. Translated by piercy M, and Berlyne D.E. Rontledge and Kegan Paul.

22 SCHOFIELD Harry, (1972), Assessment and Testing: An Introduction. George Allen and Unwin Ltd., London. ISBN 0 - 04 - 370041 - 1 Hardback.

23. SERPELL R. (1972), Intelligence Tests in the Third World. Race today, Vol. 4 No. 2, Feb., 1972 International Jounral of Psychology.

24. Psychological Service, (1973), Annual Report Government Printer, Lusaka.

25. STONES E. (1966), An Introduction to Educational Psycholoy. Richard Clay (The Chaucer Press), Ltd Bungay, Suffolk.

26. TYLER Leona, E., (1965), The Psycholoy of Human Differences; Third Edition. Meredith Publishin Co., Appleton - Century Crofts Series, University of Oregon, New York. Catalo No. 65 - 19221.

27. VALENTINE, C.W. (1963), Psychology and its Bearing on Education: Third Edition. Methuem and Co. Ltd. 36 Essex Street, Strand, WC 2. Catalog No. 2/8339/6.

28. VERNON P.E. Intellience and Cultural Environment, London, Methuem.

APPENDIX A

Pretest and posttest scores for Lake Road school.

PRETEST		POSTTEST	
SCORE X	SCORE X^2	SCORE X	SCORE X^2
83	6889	94	8836
82	6724	92	8464
81	6561	91	8281
79	6241	86	7396
76	5776	85	7225
76	5776	84	7056
75	5625	83	6889
75	5625	80	6400
73	5329	80	6400
72	5184	75	5625
71	5041	72	5184
68	4624	71	5041
68	4624	71	5041
67	4489	70	4900
66	4356	70	4900
65	4225	70	4900
64	4096	70	4900
61	3721	66	4356
61	3721	64	4096
60	3600	60	3600
$\underline{X}$ = 1423 X = 71.15	EX = 102227	$\underline{X}$ = 1534 X = 76.7	EX =119490

APPENDIX B

Scores for the hih Ability and the low Ability groups for Lake Road School.

HIGH ABILITY GROUP		LOW ABILITY GROUP	
SCORE X	SCORE X^2	SCORE X	SCORE X^2
83	6889	79	6241
82	6724	75	5625
81	6561	73	5329
76	5776	72	5184
76	5776	71	5041
75	5625	68	4624
70	4900	66	4356
67	4489	66	4356
64	4096	61	3721
61	3721	60	3600
$\underline{X}$ = 735 X = 73.5	EX = 54557	$\underline{X}$ = 691 X = 69.1	EX =48077

APPENDIX C

Pretest and posttest scores for Tree Tops school.

PRETEST		POSTTEST	
SCORE X	SCORE X^2	SCORE X	SCORE X^2
88	7744	98	9604
86	7396	92	8464
86	7396	92	8464
85	7225	90	8100
85	7225	90	8100
84	7056	90	8100
84	7056	88	7744
84	7056	85	7225
83	6889	85	7225
80	6400	84	7056
75	5625	81	6561
75	5625	81	6561
72	5184	80	6400
68	4824	74	5476
67	4489	63	3969
67	4489	63	3969
61	3721	63	3969
59	3481	45	2025
58	3364		
55	3025		
$\underline{X}$ = 1502 X = 75.1	EX = 115270	$\underline{X}$ = 1444 X = 80.22	EX =119012

APPENDIX D

Scores for the hih Ability and the low Ability groups for Tree Tops school.

HIGH ABILITY GROUP		LOW ABILITY GROUP	
SCORE X	SCORE X^2	SCORE X	SCORE X^2
88	7744	85	7225
86	7396	75	5625
86	7396	75	5625
85	7225	68	4824
84	7056	67	4489
84	7056	67	4489
84	7056	61	3721
83	6889	59	3481
80	6400	58	3364
72	5184	56	3025
$\underline{X}$ = 832 X = 83.2	EX = 69402	$\underline{X}$ = 670 X = 67.0	EX =45868

APPENDIX E

Pretest and posttest scores for Mumuni school.

PRETEST		POSTTEST	
SCORE X	SCORE X^2	SCORE X	SCORE X^2
83	6889	92	8464
80	6400	92	8464
79	6241	82	6724
78	6084	82	6724
75	5625	82	6724
73	5329	80	6400
71	5041	78	6084
67	4489	76	5776
62	3844	75	5776
62	3844	65	5625
55	3025	64	4225
47	2209	53	4096
45	2025	45	2809
41	1681	40	2025
40	1600	32	1600
29	841	28	1024
29	841	25	784
26	676	22	625
24	576		484
23	529		
$\underline{X}$ = 1089 X = 54.45	EX = 67789	$\underline{X}$ = 1097 X = 60.94	EX =75969

APPENDIX F

Scores for high Ability and the low Ability groups for Mumuni School.

HIGH ABILITY GROUP		LOW ABILITY GROUP	
SCORE X	SCORE X^2	SCORE X	SCORE X^2
83	6889	79	6241
80	6400	75	5625
79	6241	73	5329
78	6084	72	5184
75	5625	71	5041
73	5329	68	4624
71	5041	66	4356
67	4489	66	4356
62	3844	61	3721
62	3844	60	3600
$\underline{X}$ = 735 X = 73.5	EX = 53786	$\underline{X}$ = 359 X = 35.9	EX =14003

APPENDIX G

PRETEST		POSTTEST	
SCORE X	SCORE X^2	SCORE X	SCORE X^2
77	5929	78	6084
70	4900	77	5929
62	3844	68	4624
61	3721	66	4356
60	3600	65	4225
55	3025	52	2704
53	2809	47	2209
49	2401	42	1764
48	2304	30	900
38	1444	30	900
36	1296	25	625
31	961	19	361
30	900	18	324
29	841	18	324
27	729	16	256
26	676		
21	441		
20	400		
7	49		
7	49		
$\underline{X}$ = 807		$\underline{X}$ = 651	
X = 40.35	EX = 40319	X = 43.4	EX =35585

APPENDIX H

GROUP A		GROUP B	
SCORE X	SCORE X^2	SCORE X	SCORE X^2
77	5929	70	4900
62	3844	36	1296
61	3721	31	961
60	3600	30	900
55	3025	27	729
53	2809	26	676
49	2401	21	441
48	2304	20	400
38	1444	7	49
29	841	7	49
$\underline{X}$ = 532		$\underline{X}$ = 275	
X = 53.2	EX = 29918	X = 27.5	EX =10401

APPENDIX J

Upper Group and Lower Group responses to the 1986 Special Paper Two test.

UNIVERSITY OF ZAMBIA COMPUTER CENTRE — data coding sheet — Title:ITEM ANALYSIS — Programmer — Sheet 66042 — Number — Date

Sequence No.

1 5 10 15 20 30 35 40 45 50

U

43313	14123	43221	41422	42133	12112	42214	21334	44232	323
43313	14123	43222	41422	42223	12123	42313	21334	34301	223
43113	14123	43221	41422	44233	12213	42224	22334	24332	423
43410	14123	43231	41422	43233	12113	14214	21344	34222	423
43413	14123	43221	41421	42233	12113	41214	21334	43332	323
43433	14123	43241	41421	42233	12132	42311	21334	34432	433
43313	24123	43222	41422	42233	12113	42214	21334	11332	423
43413	44123	43221	41422	43233	12113	42214	21334	31334	423
43413	14123	43221	41422	43233	12113	12324	24334	34334	323
43313	14123	43211	44422	42233	12112	42214	23334	24224	121

L

21341	32413	24134	21421	32141	32432	14323	13242	13243	141
12344	22344	13132	21322	14212	21431	32421	31314	21114	432
34241	34321	24343	22123	34324	32123	43322	34343	21413	131
14233	24144	33433	24411	34421	41413	33142	34344	34443	121
14321	13223	41344	22324	12233	42142	43324	22334	41234	311
44122	33442	33342	21422	32111	14342	33142	44234	44131	412
44143	12434	13133	34223	33113	21413	12144	11234	21114	421
43123	14323	41213	32132	24132	32423	13233	42313	31214	234
41231	24123	13414	12324	12434	23141	32311	21432	12343	412

APPENDIX I

Upper Group and Lower Group responses to the 1986 Special Paper One test.

UNIVERSITY OF ZAMBIA COMPUTER CENTRE | data coding sheet | Title:ITEM ANALYSIS Programmer | Sheet 889974 Date

Sequence No.
1 5 10 15 20 25 30 35 40 45 50

U

33421	23413	14334	31421	43232	43423	43132	42133	14444	322
13421	23213	14134	31441	43332	43123	43142	42133	13433	313
33421	23213	14134	31431	43332	13123	43443	22143	13443	323
33421	23214	14134	31121	13232	43123	43142	42133	13424	322
33421	23213	14134	31421	42232	42123	43242	42141	13414	322
33421	23213	14134	31431	44232	43123	43142	42143	12142	322
33421	23113	14134	31421	43232	43123	44142	42133	13434	323
33421	23213	24134	31421	43232	43323	43042	42143	13444	323

L

44401	00000	00000	00000	00000	00000	00000	00000	00000	000
23211	22432	32413	12212	21021	04311	44211	24212	14333	341
43124	12341	32412	32214	31241	24324	31343	13211	43412	142
32413	24123	13243	14324	13142	42314	12433	24232	43241	134
14421	21221	43212	41213	41231	41134	34321	32434	11232	321
34243	42411	44233	41344	33312	31131	33322	22412	31434	234
32441	31132	13234	21423	14212	31122	42313	43412	34331	314
14341	34223	21323	32431	23142	43214	23421	42322	12234	112
33421	23213	44332	13242	13232	41143	21131	32434	21231	213
43421	21212	44133	33113	13234	32113	12123	42242	21432	414

APPENDIX K

Item difficulty and item discrimination formulae.

$$\text{Item difficulty (P)} = \frac{R}{T} \times 100$$

Where R = the number of pupils who got the item right.

$$\text{Item discrimination power} = \frac{RU - RL}{0.5 \quad T}$$

Where RU = the number of pupils in the upper group who got the item

RL = number of pupils in the lower group who got the item right.