ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

ANALYSIS OF THE DEVELOPMENT AND VALIDATION OF AN ENTRANCE EXAMINATION IN ARBAMINCH TEACHER TRAINING INSTITUTE

BY

WORKINEH TEMAMO

JUNE, 1999

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

ANALYSIS OF THE DEVELOPMENT AND VALIDATION OF AN ENTRANCE EXAMINATION IN ARBAMINCH TEACHER TRAINING INSTITUTE

BY

WORKINEH TEMAMO

JUNE, 1999

ANALYSIS OF THE DEVELOPMENT AND VALIDATION OF AN ENTRANCE EXAMINATION IN ARBAMINCH TEACHER TRAINING INSTITUTE

A THESIS PRESENTED TO THE SCHOOL
OF GRADUATE STUDIES

ADDIS ABABA UNIVERSITY

IN PARTIAL FULFILLMENT OF THE
REQUIREMENT FOR THE DEGREE
OF MASTER OF ARTS IN
EDUCATIONAL PSYCHOLOGY

BY

WORKINEH TEMAMO

JUNE, 1999

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

ANALYSIS OF THE DEVELOPMENT AND VALIDATION OF AN ENTRANCE EXAMINATION IN ARBAMINCH TEACHER TRAINING INSTITUTE

APPROVED BY EXAMINING BOARD

Habtegiorgis Berhane _____

CHAIRMAN, DEPARTMENT GRADUATE COMMITTEE

_____ SIGNATURE

Dr. AYELE MESHESHA

ADVISOR

Ayele Meshesha

SIGNATURE

Wubishet Shiferaw

EXAMINER, EXTERNAL

_____ SIGNATURE

Makonnen Yimer

EXAMINER, INTERNAL

_____ SIGNATURE

ACKNOWLEDGMENT

I am very grateful to my advisor Dr. Ayele Meshesha, Dean of the Faculty of Education and Associate Professor in the Department of Educational Psychology, for his incisive comments, continuous assistance and guidance throughout this thesis work.

Special thankful is due to Dr. Mekonnen Yimer for his helpful suggestions and comments. I also express my deepest gratitude to Dr. Habtamu Wondimu for his suggestions and comments in the initial stages of shaping the research topic.

The writer is also indebted to his instructors: Dr. Darge Wole, and Dr. Yusuf Omer, who helped him in one way or another to equip him with the necessary knowledge in various courses.

His special thanks are due to Dr. Habtegiorgis Berhane for his encouragement throughout the work of this study.

I am particularly indebted to Teshome Tesema, Tegegne Gobena, and Kiflemariam Baassa for their generous support in the collection of the data reported here.

Special acknowledgement and love goes to my wife W/ro Birhane Gizaw and to my children Eyerusalem and Gedeon Workineh who have been a source of constant encouragement, morale, material and financial support during my stay in the graduate program.

I express my appreciation to Allaho Gujo and Ethiopia, for their generous support during my stay at the University.

Finally, I also wish to thank W/ro Yeshi G/Wold who typed the thesis neatly.

TABLE OF CONTENT

LIST OF TABLES

ABSTRACT

The purpose of the present study was to evaluate the development and validation of an Entrance Examination which has been employed for selecting candidate teachers for Arbaminch Teacher Training Institute.

The subjects of the study were 583 candidate teachers who were selected from six zones of SNNPS in 1998/99 academic year.

Items of the sub-tests were analyzed for 583 subjects. Reliability coefficients were calculated for the three sub-tests by using split-half and Kuder-Richardson techniques. A correlational and regressional analysis were used for validating the three sub-tests, ESLCE GPA and High School average score. Besides, stepwise multiple regression analysis was utilized to identify variables that explain significant amount of variance in first semester GPA.

The findings of the study revealed that the sub-tests English language and mathematics have acceptable reliability. Furthermore, the study indicated that some items from each test could be revised.

Based on the findings of the present study it was concluded that the English language entrance examination, high school average score, and ESLCE GPA appeared to be better predictors of first semester GPA of the candidate teachers at the institute than mathematics and teaching aptitude entrance examination.

On the basis of the findings obtained in the study, it was recommended that since Mathematics and teaching aptitude test did not contribute significantly to the criterion variable, care should be made when constructing the items so that the reliability and predictive effectiveness of the entrance examination could be improved.

Furthermore, Utmost care needs to be taken concerning the use of the teaching aptitude test as a selection instrument in the future. If there is the need to check for general teaching aptitude in candidates, expert advice should be secured in the making of such tests.

CHAPTER ONE

1. INTRODUCTION

1.1 Background of the Problem

Even though different admissions criteria were used to select candidate teachers for Teacher Training Institutes in Ethiopia (Belay, 1990), the use of entrance examination as an admission criterion in the Teacher Training Institutes is a recent development. However, in other parts of the world academic ability tests have been used in college admissions and placements for most of this century, though their roots may be traced much further back in the history of education. They have been developed and administered by various agencies for widely different purposes with the common goal of seeking to measure a high school student's basic aptitude and preparedness for college.

Concerning admission tests, the report by Eckland (cited in Austin and Garber, 1982) revealed that modern selection testing dates back to 1900, when the College Entrance Examination Board (CEEB) was founded as a membership association by a small number of colleges, universities, and secondary schools which were concerned with the multiplicity of entrance examinations and the diversity of school curricula. Furthermore, according to Eckland the College Board introduced a new objective type of test known as Scholastic Aptitude Test (SAT) in 1926. The test was supplementary to essay tests, and was used initially in the selection of scholarship candidates from schools not preparing students for essay examinations. In addition to the College Entrance Examination Board, the American College Testing program assessment (ACT) which was founded in 1959 also widely used in the selection of applicants. Both are designed to aid college and university admission officers in evaluating the qualifications of applicants for admission.

Infact, it is known that some developed countries are mostly using standardized achievement tests to supplement teacher - made tests and to maintain the quality of educational decisions made by educators and to minimize problems associated with teacher made tests (Hoge & Butcher, 1984; Coladarci, 1986). In a country where standardized achievement tests are not in use, one has to depend largely on teacher-made tests to make different educational decisions. Even though tests are means but not ends by themselves for decision making, using entrance examination results in combination with other measures like the ESLCE GPA and High School average score help to select competent candidate teachers. It also minimizes unnecessary constraints from different angles which interfere the selection of candidate teachers as to the teaching experience of the writer of this paper in the Teacher Training Institute.

However, tests are sometimes limited in their reliability or validity because they contain items that are poorly constructed. Some items may look superficially fine but may not actually measure the construct that the test is designed to measure. One of the central problems in testing is to determine whether a specific test provides an adequate measure of a specific attribute. It is critical to determine whether or not the test meets minimal measurement requirements. Akin to this, one has to examine if test scores really reflect the attributes of the persons.

1.2 Statement of the Problem

Starting from 1996/97 Teacher Training Institutes of Southern Nations Nationalities and Peoples State (SNNPS) has been selecting candidate teachers on the basis of the combination of ESLCE GPA, High school grade average score and Entrance examination results. The entrance examination is the main interest of the present study, which is composed of three sub-tests namely, English language, Mathematics, and Teaching Aptitude.

Using reliable and valid tests for the selection of candidate teachers for Teacher Training Institutes helps to get able teachers from the training program. Many test professionals and researchers in the field of measurement and evaluation have indicated that two of the most important characterstics of a test used for selection purposes are reliability and validity. In this regard Ebel (1979) emphasized the importance of a valid test for admitting students to institutions of higher learning. The entrance examination proposal prepared by Yusuf, Azeb and Nardos (1994) for Addis Ababa University pointed out that well-designed entrance examinations, with a high level of validity and reliability, as well as carefully interpreted test scores on the basis of validity coefficients, are dependable criteria for admission.

However, some tests used for selection purposes, are limited in precision due to poor development . For instance, a validity study conducted by Alvidres and Whitworth (1981) on a locally developed college entrance examination in a Mexican University revealed that the locally developed college entrance examination lacked predictive validity due to lack of specific criteria for constructing the academic examinations and the probable lack of correspondence between test and college course content. Likewise, Meresa Abraha (1994) in his study of the Ethiopian Electric Light and power Authority Training Institute recommends that care should be taken in the development of the items of the entrance examination.

Based on this facts the current study attempts to evaluate the development and validation of a locally developed entrance examination designed to select candidate teachers for the Arbaminch Teachers Training Institute.

In general, this study attempts to answer the following basic questions:

(1) Do the items on the entrance examination differentiate candidates varying levels of ability?

(2) To what extent are the locally constructed entrance examinations reliable and valid?

(3) Does the linear combination of the various instruments i.e., the three sub-test scores, ESLCE GPA, and High School average score have a significant contribution to the prediction of academic success of the trainees?

1.3 Objectives of the Study

The purpose of this study is to investigate the development and validation of an entrance examination designed to select candidate teachers for the Arbaminch T.T.I. More specifically, the objectives are:-

(1) to identify the strengths and pitfalls of locally developed entrance examinations.

(2) to identify the validity and reliability aspect of locally prepared entrance examinations.

(3) to determine the contribution of the entrance examination results in the predication of future performance of trainees in the Institute.

1.4 Significance of the Study

The present study is mainly designed to examine critically the development and validation of a locally developed entrance examination as an admission criterion.

It is hoped that, the findings of this study might help to:-

(1) Determine whether or not the entrance examination is a valid and reliable instrument for selection of candidate teachers.

(2) Provide some suggestion about the construction of good test items.

(3) Indicate some of the problems related to the selection of candidate teachers for the Teacher Training Institute and thereby develop reliable ways for future use.

Thereby, the study may help interested individuals to conduct further research in the area.

1.5 Delimitation of the study

This study is delimited to Teacher Training Institutes of Southern Nations Nationalities and Peoples State (SNNPS) particularly to the Arbaminch Teacher Training Institute which is directed by the Southern Nations Nationalities and Peoples State Education Bureau (SNNPSEB) for its academic and administrative activity as a whole.

1.6 Definition of Important Terms

Admission Criteria: Measuring devices that are used by the Teacher Training Institutes of Southern Nations Nationalities and Peoples State.

Criterion: The external variable a test is designed to predict (Brown, 1983). First Semester GPA of candidate teachers of the Teachers Training Institute (operational definition).

Development of an Entrance Examination: refers to items of the sub-tests constructed by teachers of the institute. (operational definition)

Entrance examination: The examination prepared by teachers of the Teachers Training Institutes of SNNPS to be used as part of the admission criteria (operational definition).

Item difficulty index: The proportion of test takers in a specified group who answer an item correctly (Brown, 1983).

Item discrimination index: the degree to which a test item differentiates between people having various levels of ability or knowledge of the material tested (Brown, 1983).

Reliability: refers to the consistency or precision of evaluation results. It refers to the results obtained with an evaluation instrument and not to the instrument itself (Gronlund, 1981).

Validity: refers to the extent to which the results of an evaluation procedure serve the particular uses for which they are intended. It refers to the results of a test or evaluation instrument for a given group of individuals not to the instrument itself (Gronlund, 1981).

Predictive validity: refers to the predictive power of the predictor variables to forecast the success of a candidate teacher at the end of the semester or at the end of the training program. (operational definition)

CHAPTER TWO

2. REVIEW OF RELATED LITERATURE

In this section different studies carried out by different researchers in the area in Ethiopia and in other countries, which are directly and tangentially related to the present study, were reviewed under the following main headings:

2.1 Tests in selection and admission.

2.2 Evaluation of the development and validation of tests.

2.3 An overview of the admission criteria.

2.1 Tests in Selection and Admission

2.1.1 Uses of Tests in Selection Decision

Tests are used for many purposes and test scores help for a variety of decision making. The usefulness of a test will depend very much on the type of decisions one wishes to make. An institution or organization may decide that some persons are acceptable while others may not; those not acceptable are rejected and no longer become the concern of the institution or organization. This feature - rejection and the elimination of those rejected from the immediate concern of the institution, is central to a selection decision (Cronbach and Gleser cited in Nitiko, 1983). According to Sax (1974) selection decisions demand of tests the ability to predict success and failure with minimum risks both to the institution and to the individuals involved.

Furthermore, as Chauncey and Dobbin (1963) stated, the selection should be done in such a way as to assure the best possible education for the greatest possible number of candidates; meaning that the institution should discover those applicants who are best able to profit from the teaching that it has to offer. Moreover, selection decisions are greatly improved when a close

correspondence exists between the testing tasks and those skills, abilities, and attitudes required for success in the institution.

Even though tests are used to make important decisions about individuals, unfortunately, one feature that all tests share in common is their limited accuracy. They rarely, if ever, provide exact, and definitive measures of variables that are believed to have important effects on human traits. Studies shows that tests do not provide a basis for making completely accurate decisions about individuals. In reality there is no method that guarantees complete accuracy. Although tests are far from perfect, they generally give the best, fairest, and most economic estimate of obtaining the information necessary to make sensible decisions about individuals (Wigdor and Garner, 1982). Maintaining this opinion, Gronlund (1981) stressed that to argue that better educational decisions would be made without test scores is to argue that better decisions are made when less information is available.

Following the above rationale, recently different institutions in Ethiopia use tests for selection of applicants to their training programs (Kebede, 1991; Tamiru, 1992; Meresa, 1994). According to the Testing Center technical report by Ayele Meshesha and Langmuir, (1971) some institutions have been using tests for assessment and selection.

2.1.2 Admission Tests

Admission tests can be used to help an institution to assess the information it receives from high schools about students applying for admission. The performance of an applicant on a test has some meaning that is not contained in any of the other records about him. His test score is the one piece of evidence about him that is most nearly comparable with similar evidence about all other candidates. The test is a common denominator that permits the college to make allowances for differences among schools and communities.

Supporting the above notion, Chauncey et al., (1963) pointed out that the addition of test information gives the admissions officer or the institution a different source of insight into the learning characteristics of the applicant, insight which he could not have applied to the admissions decision on the basis of the transcript alone. And the admissions officer uses the test performance information with the high school record in reaching his decision about each applicant.

2.2 Evaluation of the Development and Validation of Tests

When tests are used for selection, it is imperative to show that the scores on these tests bear a relationship to success in the program for which the institution or organization has selected able persons. If investigation does not show these tests can distinguish between those likely to succeed and those that don't, such tests should be improved upon or eliminated.

Selection decisions need not be perfect, and tests cannot be expected to have perfect validity for selection. Moreover the use of tests in selection results in that some applicants would have been successful had they been selected instead of rejected, and some, even though they were accepted, may turn out to be unsuccessful. Hence, the tests are evaluated in terms of their consequences, which is the main concern of the present study.

2.2.1 Evaluation of the development of Tests

The construction of valid and reliable tests requires consideration of sound application of established principles of test construction and quantitative information regarding the difficulty and discrimination power of each test item. In this, regard Engelhart (1975) pointed out that the production of a valid test item is both an art and a science. The scientific aspects involve technical knowledge and skills regarding item structure and format, discrimination, difficulty, and reliability. Art comes into play in integrating the scientific, the verbal, and

practical aspects of the total operation. Furthermore, Cohen, Swerdlik, and Smith (1992) stated that all tests are not created equal. The creation of a "good test" is not a matter of chance; it is the product of the thoughtful and sound application of established principles of test construction.

Regarding the above opinions, numerous studies have been conducted on the effects of violating selected item construction principles, item arrangement and test performance, and item difficulty and discrimination. For instance, Chase (1964) estimated the word count ratio at which extra length correct options in multiple choice items would induce a set to select these options, and then showed that practice on short option, easy items would alter this set. The study of Terranova (cited in Mcmorris, Brown, and Pruzek, 1975) reveled that items with positive stems have been found to be easier than those with negative stems. Further, items were easier when there were a minimum number of changes in stem modes, but reliability seemed unaffected.

Furthermore, the study of Dun and Goldstein (cited in Mcmorris et.al., 1975) reveals that violating item construction principles makes no difference on validity and reliability, but the faulted items were easier. Likewise, Mcmorris, et al., (1975) pointed out that item-writing faults are neither extremely potent, nor nil in their effects on item scores, but conventional validity and reliability coefficients may tend not to be much affected by these types of faults. Moreover, Williamson and Hopkins (1975) compared the effects on validity and reliability of a "none-of-these" option versus a homogeneous fifth option, and found no significant differences.

Concerning item arrangement and test performance, different studies reveal that item arrangement significantly influences test performances, but this relationship appeared when the examinations were administered as speed tests rather than power tests (Brenner; Sax and Cromack cited in Marso, 1970). The experimental study conducted by Marso (1970) suggests that, even on relatively lengthy examinations which are administered as classroom or power test, one

need not be greatly concerned with arranging test items by difficulty, in order of class presentation or by similarity of content covered at least in terms of influencing required testing time or student performance. However, on a logical versus empirical basis, one might still advise a teacher to place a few easy items at the beginning of a test to possibly ease initial test anxiety; but any additional time spent on arranging items by difficulty or content might more profitably be spent on improving teacher-devised tests in other ways.

When evaluating the development of a test, the most important aspect to be evaluated is the quality of the items. Close examination of individual items is essential to understand why a test shows specific levels of reliability and validity. In this regard Anastasi (1988) pointed out that both validity and reliability of any test depend on ultimately on the characteristics of its items. Furthermore, Brown (1983) stated that a test is as good as the items it contains.

The influence of item difficulty and discrimination on the properties of the whole test was seen in the study of Cronbach and Warrington; Gulliksen; Lord; Ray and Hundleby and Goldstein (cited in Henrysson, 1971). The findings of their study revealed that test reliability is maximized when items are at the same difficulty level, and the level should be near 50% as possible. Furthermore, the general agreement of test professionals (Myers, 1962) indicated that for maximum reliability, tests should be homogeneous with respect to item difficulty. Validity is usually considered to be of greater importance and reliability is often sought after mainly as a means to an end of increased validity. It is generally accepted that, other things being equal, an increase in reliability will produce an increase in validity although the increase in validity will be smaller (Brogden; Ferguson; and Tucker; cited in Myers, 1962).

When a test is to be used for dividing a group of persons being processed into accepted and rejected candidates, a test which will discriminate the poorest acceptable applicant from the best of the rejected group is desired. In this regard, Henrysson (1971) pointed out that for tests intended for selection, the

bulk of items must be at difficulty level appropriate for those at the ability level where selection will be made. Supporting this opinion, different studies reveal that if items are chosen for maximum discrimination, the items will tend to have difficulty values in the vicinity of 50 percent (Ebel and Damrin, 1958; Anastasi, 1976; Cronbach and Warrington, 1952). Some studies even indicated that if the goal of measurement is for selection (to maximize differences between groups), items with difficulties of .50 to .70 are best since there is little difference in discrimination power between items with difficulties of .50 and .70 (Brown, 1983).

2.2.2 Validation of Tests: Criterion-related Validity

Various procedures have been utilized in determining test validity. But experts have attempted to categorize them into three major components (i.e. content validity, criterion-related validity, and construct validity). Among these classifications criterion-related validity is another major concern of the present study.

The appropriate method of validating a test will depend on the use of the test. When tests are used for decision-making, evidence of criterion-related validity is needed to determine whether the test adequately performs its intended function (Brown, 1983; Murphy and Davidshofer, 1994).

In general terms, criterion - related validity studies pertain to the techniques of inspecting the empirical relationship between test scores and certain criterion measure(s). In this case a test user determines the extent to which performance on the test fortells future achievement either on the job or training (i.e. predictive validity). Accordingly, criterion related validity is evaluated interms of the role and contribution of a test to increase predictive or decision-making accuracy (Anastasi, 1976; Nitiko, 1983; Brown, 1983; Wiersma and Jurs, 1985).

2.2.2.1 *Validity of Admission Criteria for Selection Decision*

When the purpose of admission tests is to forecast future endeavor, say, success in learning, training, or on the job, the tests act as predictors, and the variable predicted as criterion measure. In addition, it can be argued that if test scores predict with fair accuracy students future performance on a task, one can be sure that the test has predictive validity for the purpose.

Concerning the above issue, Anastasi (1976) claimed that any information provided by prediction studies is highly important to tests used in the selection of students. That is, selecting students for admitting to colleges, universities or professional schools are instances of certain decisions requiring knowledge of the predictive validity of tests.

Maintaining the ideas so far mentioned, Brown, (1983); Nitiko (1983), and Weirsma and Jurs (1985) indicated that in any selection situation, say admission to colleges, training programs for scholarship competitions, and hiring for jobs, tests, along with other measures, are used to get the necessary information. Accordingly, it is common knowledge that teacher-made tests, school-oriented tests, published tests, school grades, and biographical information are common variables that are used as predictors of educational outcomes.

Hence, predictive validity is particularly important in tests used for selection decisions such as in admission to Teacher Training Institutions and other training programs. The scores earned by candidate teachers on pre-admission variables are compared with first or second semester grades earned in the Teacher Training Institute to examine if the pre-admission variables have really helped in selecting the right individuals. However, predictive validity data are affected by some factors such as lack of reliability of tests; lack of uniform grading system; restricted range, and lack of reliability of criterion measures (Gronlund, 1981; Cohen et.al., 1992; Michael, 1965; Brown 1983). In this connection, Guilford (1965) stated that the difficulty of obtaining adequate

criteria of what to be measured is an indication of the problem for the variation of validity coefficient. Furthermore, in discussing validity index, he suggested that a writer should be careful to state all the conditions related to the size of the index obtained, and any reader should accept interpretations only when the significant circumstances are kept in mind.

2.2.2.2 Choice of Criterion Measure for Validation

Evidently, criterion-related validity studies focus on assessing the extent to which scores on a test permit judgement about an individuals likely standing on a criterion measure. Nevertheless, ones judgement on the adequacy of the predictive power of a test depends also on the appropriateness of the criterion measure used. It is well-known that obtaining adequate criteria is one of the problems encountered in validity studies. In spite of this short coming most studies, usually use academic achievement as practical criteria for determining predictive validity.

Supporting the above notion, Anastasi (1976) also pointed out that school grades, test scores, promotion and graduation records, special honors and awards and teachers' ratings are the criteria most frequently employed in validating intelligence tests or some index of academic achievement. Even though a variety of criterion measures are employed for different purposes, most of the time they tend to be achievement test scores, ratings, marks and other quantified judgements in the field of education (Brown, 1985; Nitiko, 1983).

Hence, it can be said that many sources indicate that various indices of academic achievement provide data as criterion measures at all educational levels from primary grades to higher learning institutions.

2.2.3 Validity Studies on Selection Instruments

As it is mentioned earlier, Teacher Training Institutes of SNNPS have been selecting candidate teachers on the basis of a combination of Entrance examination results; ESLCE GPA, and High School grade average point. In this part related studies to the selection instruments will be discussed.

2.2.3.1 Entrance Examination

Evidence in defense of ability tests in college admissions come largely from validity studies of such tests in predicting educational attainment, especially the academic performance of students in institutions of higher learning. Austin and Garber (1982) pointed out that admissions tests have predictive power in college, independent of one's high school grades, partly because grades depend considerably on the academic and competitive environment of a particular high school a student attended.

Several studies done in Ethiopia revealed that entrance examination is the best predictor of academic performance in both vocational and professional training institutes. For instance, a validity study conducted by Tamiru (1992) in Catering and Tourism Commission Institute found that entrance examination is the best potential predictor of the performance of the candidates trained at the institute. The study found a correlation coefficient of 0.84 between entrance examination and the cumulative average scores of the practical training performance and an r= 0.77 between entrance examination and the cumulative average scores of the training performance GPA in food and Beverage control department.

Furthermore, several authors (Lakew, 1972; Mittman, 1972; Bowers, 1968; Mekonnen et.al., 1991) explained that aptitude tests might be considered as the most important predictors in the case of academic performance in college.

For example, Mekonnen et.al., (1991) reported a moderate correlation (r = 0.53) between the locally designed aptitude test scores and freshman grade.

However, the findings of some studies show slightly lower relationship between entrance examination results and College GPA. For instance, the study of Meresa (1994) in the Ethiopian Electric Light and Power Authority Training Institute revealed that even though entrance examination results predict better than ESLCE GPA and high school grade average score, the relationship between entrance examination results and GPA was low. The study found a correlation coefficient of .111 between English entrance examination and first semester average; also a correlation coefficient of .202 between Math's entrance examination and first semester average, and a correlation coefficient of .197 between aptitude test and first semester average.

In other instance, the validity study conducted on locally developed entrance examination in a Mexican University (Alvidres and Whitworth, 1981) revealed that locally developed entrance examination lacks predictive validity. The study found a correlation coefficient of .36 between mathematics entrance examination and university GPA and a correlation coefficient of .26 between English entrance examination and University GPA.

The literature shows both the strength and weaknesses of entrance examination in predicting success. Despite this, entrance examinations are used allover the world for the same purpose. The ones in use in the Teacher Training Institutes of SNNPS are meant to do just that. The present study will examine and see if they achieve their objective.

2.2.3.2 Ethiopian School Leaving Certificate Examination (ESLCE)

The Ethiopian School Leaving Certificate Examination has been used by different people and organizations for different purposes. Higher education

institutions have been using it for the selection and admission of candidates to their units.

Different individuals who have studied about the predictive validity of the ESLCE stated that the ESLCE contribute for the predictive quality of college performance (King and King, 1972; Bowers, 1969). Some studies also revealed that the ESLCE was one of the predictors of college success (Bowers, 1968; King and King, 1972). A predictive validity study conducted on Teacher's Training Institutes in Ethiopia (Belay, 1990) reveals that ESLCE GPA has significant and positive correlation with Teachers' Training Institute grades on language, social science, aesthetic, and natural science courses.

Despite the strengths reported above, the findings of many other studies have indicated that ESLCE GPA is inadequate for predicting college performance of students (Mekonnen et.al., 1991). The overall ESLCE GPA and some grades of individual ESLCE subjects have also a relatively weak relationship with college academic performance (Melaku, 1975; Madsen, 1969; Mittman, 1972). However, Tracy (1965) indicated that the correlation between the overall ESLCE GPA and first year cumulative GPA (r=0.36) was slightly better than the relationship between individual ESLCE subjects and freshman grades.

Because of the weaknesses of the ESLCE in predicting college success, some studies suggest the pressing need for additional admission criteria (Mekonnen, et.al., 1991; Tirusew, Getachew and Desalegn, 1992). Furthermore, the study of Meresa (1994) in the Ethiopian Electric Light and Power Authority Institution recommended that care should be made regarding the usefulness of ESLCE GPA in the selection of candidates to the training Institute.

Even though, the above studies suggest the weakness of ESLCE in predicting the success, different institutions still continue to use it as an admission criterion to their training programs.

2.2.3.3 High School Grade Average (HSGA)

Different institutions in Ethiopia use high school grade average (HSGA) as an admission criterion. (Belay, 1990; Kebede, 1991; Meresa, 1994;. Tamiru, 1992; Shenkute, 1991). The present study also tries to examine the predictive power of high school grade average which is in use in the Teacher Training Institutes of SNNPS. It is used in combination with others as admission criteria.

Different studies conducted in other countries to examine the predictive validity of High School grade average reveal that HSGA is a more valid predictor of college success (Franz, Daviz, and Garcia, 1958; Mann, 1961; Michael and Jones, 1963; Spaulding, 1959; Chissom and Lanier, 1975). Among studies concerned with grade predictions in specific subject areas, Passons (1967) concluded, that, high school achievement was the most predictive indicator of future college success. Michael et.al., (1962) concluded that high school grade average is a valid predictor of success for both sexes in college.

Furthermore, Coppedge (1969) concluded that colleges are justified in using such variables as high school grades for college admission purposes. The correlation between high school grades average and college success seemed to indicate that the major objectives of higher education, as demonstrated by the curriculum and expected standards of achievement, are not too dissimilar to those of the public secondary schools.

Despite the strengths reported above, some studies reveal that, high school grade average (HSGA) is less accurate as a predictor of subsequent academic performance (Ramist, in Yusuf et.al.,1994). The reasons behind is that, grading standards differ from one secondary school to another. Grades do not represent common matrix; and due to "grade inflation", the range of mean secondary school grades has become narrow. Furthermore, some studies found a reduction in the predictive validity when forecasting college freshman grade

point average from high school achievement data (Michael et.al.,1963; Dalton, 1976).

2.2.3.4 Combination of Variables in Predicting Academic Success

Several studies combine various independent variables together to predict the success of students in the institutions of higher learning. It has usually been suggested that multiple correlation coefficients have been computed by many investigators to discover the extent to which a combination of predictive factors would improve the correlation with the criterion over that given by a single predictor variable. In the study conducted by Mekonnen et.al., (1991) a relatively significant multiple correlation coefficient was reported when the combination of aptitude test score, high school average score and ESLCE GPA were used in the prediction of freshman GPA. The multiple correlation coefficient was found to be 0.49.

Meresa (1994) also reported statistically significant multiple correlation coefficient of the composite of seven predictors with first year GPA of the trainees. The multiple correlation coefficient was found to be 0.30. Similarly, a substantial result was reported by Jacobson (1963) when the combination between high school average and scholastic aptitude test was used. The correlation coefficient with the established criterion measure was found to be 0.55.

In his study Mann (1961) also reported a very substantial multiple correlation coefficient in the prediction of college academic success. In the combination of SAT both verbal and mathematical with high school average, he found a multiple correlation of 0.82. Willingham (1964) also found a substantial relationship between college GPA and the combination of high school average score and SAT mathematics and verbal score. The multiple correlation coefficient established was 0.60. Moreover, Chissom and Lanier (1987) also reported a significant overall correlation coefficient of 0.57 when the combination

of high school GPA, SAT-verbal, SAT-Mathematical was made for the prediction of college success.

In general, studies seem to indicate that better predication of success in higher education can be made when composites of several predictor variables are used in certain linear combination.

Based on this fact, the present study also tries to see the contribution of the combination of predictor variables on the success of trainees in the Teacher Training Institute.

2.3 An Overview of the Admission Criteria

Different studies have been done to examine some of the most commonly used criteria by school districts in other countries to recruit teacher candidates. Sinclair and Picogna (cited in Applegate, 1987) surveyed state directors of teacher preparing institutions in America that used selective admission criteria to select teacher candidates. From the institutional report they have identified six categories of selection techniques: "academic grade point average; desire or commitment to teaching; satisfactory physical and mental health; voice quality and appearance; satisfactory completion of a course in oral communication or speech; and satisfactory personality or character".

Different studies suggest that an admission criteria used determine the quality of the trainees future performance. Thorndike (1981) stated that the purpose of a system for selection decision is to hire or admit those individuals who will maximize the total value or utility to institutions.

Haberman (1974) summarized the following guidelines in the process of teacher candidate selection:

1. Admission to professional education is a professional decision not a student right.

2. Selection Criteria should drive from the program goals and the capabilities needed by the individuals to achieve these goals.

3. Selection is a process, not an event. More rigid adherence to existing criteria will not improve selection.

4. Selection must assess the potential of candidates to function as continuous learners.

In the history of Ethiopian teacher education, different criteria have been used for admitting teacher candidates for primary teachers' training. According to one report from the Ministry of Education (1972-1989 E.C.) the following were put in practice.

- From 1972 - 1980 students completing secondary school with a class rank of 20 or less were considered for admission.

- From 1981 - 1985, high school average academic performance, ESLCE GPA, interviews, teaching experience and a letter of recommendation about the conduct and community participation of the candidate were used as elements of the admission criteria.

- From 1987 - 1989, ESLCE GPA, interviews and recommendations were used as selection criteria. However, applicants participating in the National Military service, after completing grade 12 were admitted without considering the above.

- Recently, most teacher training institutions of regional governments use a combination of entrance examination results; ESLCE GPA and High School grade average score as a measuring device for admission which are the issues of the present study.

Next, a description of the methodology used in the study will be given.

CHAPTER THREE

3. METHODOLOGY

3.1 Entrance Examination

The entrance examination used in this study was constructed locally by the teachers of Arbaminch Teacher Training Institute. This examination was an assessment of academic ability and teaching aptitude. It consisted of the following three sub-tests:

1. **English Language:** the questions comprising the first sub-test were written by teachers who taught English at the institute. The test consists of 40 items and was scored from 0 to 40 and converted to make 10 percent of the total score.
2. **Mathematics:** the items in the second sub-test were written by mathematics teacher of the institute. The test consists of 30 items and was scored from 0 to 30 and converted to make 10% of the total score.
3. **Teaching Aptitude:** the third sub-test was prepared by a member of the institute who used no specific criteria other than his own knowledge of aptitude tests. The test consists of 30 items and was scored from 0 to 30 and converted to make 20% of the total score.

So, the entrance examination on the three sub-tests made 40 percent of the admission criteria. The remaining 60 percent was allotted to the ESLCE GPA (35%) and high school average score (25%).

3.2 Subjects

The subjects consisted of the entire candidates who entered the Arbaminch Teacher Training Institute in the 1998/99 academic year. There was a total of 590 candidate teachers of whom 395 were males and 195 were

females. Seven candidates who had incomplete data were excluded from the sample and the remaining 583-candidate teachers were included in the study.

3.3 Procedures of Data Collection

Answer sheets of the three sub-tests of the 583 candidate teachers which had been corrected by selected scorers and the results of the sub-tests were collected from the Southern Nations Nationalities and Peoples State Education Bureau (SNNPSEB). ESLCE GPA, High School average score, and first Semester grade point average of the candidate teachers of the Institute were collected from the record office of the Arbaminch Teachers Training Institute.

3.4 Variables

The present study tries to evaluate the development of items and the predictive power of the entrance examination. To evaluate the development of this entrance examination, critical analysis were carried out.

To evaluate the predictive power of an admission criteria, the following variables are included in the study and their representations are as follows:

3.4.1 Independent Variables

X_1 = Entrance examination results in English Language.
X_2 = Entrance examination results in Mathematics.
X_3 = Entrance examination results in Teaching Aptitude.
X_4 = ESLCE GPA score.
X_5 = High School Average Score.

3.4.2 Dependent Variable

Y = The candidate teachers first semester GPA.

3.5 Procedures of Data Analysis

Entrance examination answer sheets, were ranked in descending order (from 1 to 583) and dichotomized 0 for wrong and 1 for correct responses. Item analysis was conducted to determine difficulty and discrimination indices. Spearman- Brown formula and Kuder-Richardson formula were conducted to determine the reliability coefficient of the entrance examination by using the item analysis data.

Descriptive statistics and intercorrelations among all predictor and criterion variables were calculated for the total sample of candidate teachers. Multiple regression and a stepwise multiple regression analysis involving the 5 predictors and one criterion variable were conducted.

CHAPTER FOUR

4. RESULTS AND DISCUSSION

The results of the study have been summarized under the following subheadings: results of the item-analysis data; results of the correlational analysis; results of a multiple regression analysis and results of the stepwise regression analysis. Relevant or accepted procedures and guidelines important in evaluating and interpreting the results of the study are also presented.

4.1 Results

4.1.1 Results obtained using item-analysis Data

Several studies, on the evaluation of the discrimination of items, suggest the optimal boundary lines to demarcate the upper and lower areas of a distribution of scores. For instance, Henrysson (1971) suggests taking the upper and lower one-third of examinees. D'Agostino and Cureton (1975) contend that 21 percent is slightly better. Cureton(1975) revealed that as the distribution of test scores becomes more platykurtic, the optimal boundary line for defining upper and lower gets larger and approaches 33 percent. The study of Kelley (cited in Cohen et.al., 1992) revealed that upper and lower 27 percent of the group is an accepted practice in the evaluation of the discrimination of items scored 0 or 1 provided that the distribution is normal. Furthermore, Feldt (1963) argued that 27 percent yield highly efficient estimates and thus recommended it as the better method in item analysis procedure. Moreover, Allen and Yen (cited in Cohen et.al., 1992) pointed out that for most applications, any percentage between 25 and 33 will yield similar estimates. Since 27 percent is efficient and is an accepted practice by different researchers (Ebel, 1975; Feldt, 1963), the

present study used the 27 percent to demarcate the upper and lower areas of the distribution of scores of the entrance examination.

Furthermore, different test specialists suggested guidelines which are important and relevant to evaluate items of the entrance examination. In this regard Ebel (1979) suggested the following guidelines which are relevant only for describing the contribution of an item to a test's reliability.

Index of Discrimination	**Item Evaluation**
.40 and up	very good item
.30 -.39	good item
.20 - .29	reasonably good item
.10 - .19	Marginal item, usually subject to improvement
Below .10	poor item, to be rejected or revised.

Moreover, Ebel (1966) suggested the following guidelines for interpreting item discrimination index values.

Index of Discrimination	**Item Evaluation**
.40 and up	high
.20 -.40	Moderate
.01 - .20	Low
Zero or Negative	Poor item, to be rejected

The suggestions regarding the discrimination power of an item given above has been used in the study.

In order to evaluate and interpret the reliability coefficient of the entrance examination, the following guidelines are used in the present study which are suggested by Murphy and David Shofer, (1994). For almost any testing application:-

1. Reliability estimates of .90 or more regarded as a high level of reliability
2. Reliability estimates of .80 or more are typically regarded as moderate to high
3. Reliability estimates around .70 are to be regarded, as low, this level of reliability might be sufficient when tests are used for screening or for making preliminary decisions.
4. Reliability estimates lower than .60 usually are thought to indicate unacceptably low levels of reliability. (See Murphy and David Shofer, 1994, pp. 103-104 for more justification).

Based on the guidelines suggested above, results obtained by the item analysis procedures are interpreted as follows.

As it is seen from the item analysis result of the English language entrance examination (Table 1) , 7 items (17.5%) discriminate highly and 22 items (55%) discriminate moderately. They are reasonably good items and hence, they contribute better to the reliability of the entrance examination.

Whereas 10 items (25%) of the entrance examination have low discrimination index. Such items may not contribute to the reliability of the entrance examination or they contribute little. Furthermore one item (2.5%) does not discriminate at all. (see appendix D). It is probably miskeyed, or is at least seriously ambiguous. In its present form it lowers the reliability of the entrance examination, and validity as well.

Furthermore, as observed from results, the English language entrance examination had a split-half reliability of 0.70 and internal consistency (KR_{20}) of 0.70 with a mean of 27.07 and standard deviation of 4.69.

Table 1. Test analysis report (See Appendix D)

Test Title :- English language Entrance Examination.

Number of examinees:- 583

Characteristic

Discrimination

A.	**Item**	
1.	High (.41 and up) -	17.5%
2.	Moderate (.21 to .40) -	55%
3.	Low (.01 to .20) -	25%
4.	Zero or Negative -	2.5%
B	**Score**	
1.	Mean -	27.07
2.	Standard deviation -	4.69
3.	Reliability	
	a) Split-half -	0.70
	b) KR_{20} -	0.70

With regard to the mathematics entrance examination (Table 2), the results of item analysis data revealed that 11 items (36.67%) had a high discrimination power and a similar number discriminate moderately. Both groups of items which discriminate highly and moderately contribute well to the reliability as well as to the validity of the entrance examination. Where as, 6 items(20%) have low discrimination power and 2 items (6.67%) discriminate negatively. In general, some 26.67% of the items do not contribute well to the reliability of the entrance examination. As observed from the result (Table 2), Mathematics entrance examination had split-half reliability of 0.69 and KR-20 of 0.66.

Table 2. Test Analysis Report. (See Appendix E)

Test Title: Mathematics Entrance Examination.

Number of examinees: 583.

Characteristic

Discrimination

A. Item

1. High (.41 and up) - 36.67%
2. Moderate (.21 to .40) - 36.67%
3. Low (.01 to .20) - 20%
4. Zero or Negative - 6.67%

B. Score

1. Mean - 15.187
2. Standard deviation - 4.123
3. Reliability
   a) Split-half - 0.69
   b) KR_{20} - 0.66

In a similar manner, the results of item analysis data of the sub-test of the entrance examination, teaching aptitude (Table 3) revealed that from the total of 30 items, 4 items only (13.33%) discriminate the test takers varying levels of ability, highly and 15 items (50%) discriminate moderately. These results suggest that the items are either very good items or reasonably good items which have a great contribution to the reliability as well as to the validity of the entrance examination. Whereas one item (3.33%) discriminates negatively and in the present form this item is intrinsically ambiguous and it reduces the reliability of the entrance examination. The remaining 10 items (33.33%) have a low discrimination power. Hence they contribute low to the reliability as well as to the validity of the entrance examination.

Furthermore, as it is observed from the results of test analysis report (Table 3) teaching aptitude entrance examination had a split-half reliability coefficient of 0.60 and internal consistency (KR_{20}) reliability coefficient of 0.63 with a mean of 20.44 and standard deviation of 3.47.

Table 3. Test Analysis Report. (See Appendix F)

Test Title: Teaching Aptitude Entrance Examination.
Number of Examinees: 583.

Characteristic

Discrimination

A. Item

1. High (.41 and up) - 13.33%
2. Moderate (.21 to .40) - 50%
3. Low (.01 to .20) - 33.33%
4. Zero or Negative - 3.33%

B. Score

1. Mean - 20.44
2. Standard Deviation - 3.47
3. Reliability
   a) Split-half - 0.60
   b) KR_{20} - 0.63

4.1.2 Results obtained using Correlational Analysis

A key index in predictive validity is the correlation coefficient which indicates the degree of relationship that exists between the scores individuals obtain on two instruments – the predictor and criterion.

Since the interpretation of validity coefficient must take into account a number of concomitant circumstances, there is no general answer to the question "How high should a validity coefficient be?". However, some measurement specialists such as Brown (1983) and Anastasi (1988) suggested that the obtained correlation should be high enough to predict the criterion from test scores. According to Cronbach (1990), the only sensible answer to the question "What is a good validity coefficient?" is the best one can get. He stressed that a test may predict efficiently if it shows any significant correlation with the criterion, however low.

Regarding the question "How high should a validity coefficient be?" Fraenkel and Wallen (1993) pointed out that correlation coefficients below 0.35 show only a slight relationship between variables. Such relationships have almost no value in any predictive sense. According to Fraenkel et.al., (1993) to check validity of scores, the correlation coefficient should be at least 0.50, and preferably higher.

Other scholars like Cohen and Manion (1994) suggested three different approaches which are helpful in interpreting the magnitude of the relationship between predictor variables and Criterion measures; namely: examining the strength of the relationship; examining the statistical significance of the relationship and examining the square of the correlation coefficient. Usually these techniques are employed whenever an attempt is made to explain the nature and degree of relationship between independent variables and criterion measures. Furthermore, they suggested the following guidelines which are helpful in evaluating the correlation coefficient obtained.

1. Correlation Coefficients from 0.20 to 0.35 show very slight relationship between vairables even if they are statistically significant.

2. Correlation Coefficients ranging from 0.35 to 0.65 inclusive are considered as helpful for group prediction and are useful when combined with other correlation in a multiple regression equation.

3. Correlation Coefficients ranging from 0.65 to 0.85 inclusive are considered accurate for prediction purposes.

4. Correlation Coefficients over 0.85 are considered to show a close relationship between the two variables correlated.

In order to evaluate the Correlation Coefficient obtained in the present study, the suggestions of Cohen et.al., (1994) are used as guidelines since they go in line with the suggestions of others mentioned above.

The results of the intercorrelations among the five independent variables and the first Semester GPA, the dependent variable along with the means and standard deviations for each measure are shown in Table 4. As may be observed, one sub-test of the entrance examination, English language (X_1) had the highest correlation (r=.47) with first Semester GPA accounting for 22.5 percent of the total variance in the criterion measure. The other sub-tests of the entrance examination Mathematics (X_2) and Teaching aptitude (X_3) have slight relationship with the criterion measure (r=0.32 and r=0.35 respectively). The explained variance accounted by mathematics entrance examination is slightly lower (r^2= 10.0 percent) than teaching aptitude (r^2= 12.2 percent). The three sub-tests also some what highly intercorrelated with each other (r=0.38 between English and mathematics, r=0.54 between English and teaching aptitude and r=0.28 between Mathematics and teaching aptitude) and so added little to the overall predictability.

Besides the three sub-tests, the independent variables ESLCE GPA (X_4) and high school average (X_5) accounted for 14.4 percent (r=0.379) and 38.2 percent (r=0.626) of the total variance in the criterion measure respectively.

Close examination of Table 4 indicates that all the predictor variables have positive relationship with the criterion measure (y) and the degree of

relationship is relatively low. Furthermore, the predictor variables which have high standard deviations show a greater spread or variation on the performance of the Candidates than those whose standard deviation is relatively low. The Criterion variable has a low standard deviation when compared with predictor variables. This indicates that there is a low variation on the performance of the candidates.

Table 4. Intercorrelations, Means, Standard Deviations and Proportion of variance to the prediction of First Semester GPA .

Variable	x_1	x_2	x_3	x_4	x_5	Y	$\bar{x}$	S	r^2	F
x_1	1.000	0.383*	0.535*	0.420*	0.492*	0.474*	6.446	1.205	0.225	168.686*
x_2		1.000	0.278*	0.314*	0.341*	0.316*	4.812	1.420	0.100	64.363*
X_3			1.000	0.356*	0.414*	0.349*	13.198	2.471	0.122	80.732*
x_4				1.000	0.435*	0.379*	17.813	3.155	0.144	97.245*
x_5					1.000	0.626*	13.110	2.679	0.392	347.852*
y						1.000	2.495	0.563		

*Significant at 0.05 level

4.1.3 Results of Multiple Regression Analysis

As indicated in Table 5, below the three sub-test scores together contributed 25.6 percent to the explained variance of the first Semester GPA. Interms of the magnitude of the regression coefficient, the important independent variable in descending order are X_1, X_2, and X_3. The F-value of = 66.446 indicates that the contribution of the three sub-tests of the entrance examination for predicting first semester GPA is statistically significant at the .05. In addition, the result indicated that English language (X_1) is more important than mathematics (X_2) and teaching aptitude (X_3) in predicting candidate teachers academic performance.

Table 5 . Results of Multiple Regression Analysis of the three sub-test scores on first semester GPA

Variable	Regression Coefficient	Standard error of regression Coefficient	Regression Weight (β)	T-values
X_1	0.166	0.021	0.354	8.000*
X_2	0.058	0.015	0.147	3.775*
X_3	0.027	0.010	0.119	2.789*

* Significant at .05 level.

Multiple Correlation = 0.506 (R^2=0.256)

Constant = 0.756

F-value = 66.446

Standard error of estimate = 0.487

As it can be seen from Table 6, below the five predictors together contributed 43.6 percent to the explained variance of the first semester GPA. In terms of the magnitude of the standardized regression coefficient, the important variables in descending order are X_5, X_1, X_2, X_4 and X_3. The F-value of 89.316 indicated that the contribution of all independent variables for predicting the first semester GPA is statistically significant at the .05 level. Furthermore, the T-values of X_1, X_4 and X_5 shows that these variables are highly significant to the prediction of first semester GPA. However, the T-values of X_2 and X_3 show that these variables do not contribute significantly to the prediction of first semester GPA.

Table 6. Results of Multiple Regression Analysis of the five predictor Variables with first semester GPA

Variable	Regression Coefficient	Standard error of regression coefficient	Regression weight (β)	T	P
x_1	0.083	0.019	0.191	4.433*	0.000
x_2	0.023	0.014	0.072	1.631	0.103
x_3	0.003	0.009	0.001	0.346	0.729
x_4	0.013	0.006	0.056	1.956*	0.051
x_5	0.101	0.008	0.486	12.573*	0.000

*Significant at .05 level

Multiple Correlation = 0.661 (R^2=0.436)
Constant = 0.237
F-value = 89.316
Standard error of estimate = 0.425

Regression equation

$y' = 0.237 + 0.083x_1 + 0.023x_2 + 0.003x_3 + 0.013x_4 + 0.101x_5$

4.1.4 . The Stepwise Multiple Regression Analysis

The stepwise regression method has been employed in this study to evaluate the relative contribution of each predictor variable in forecasting the criterion measure and to identify the potential predictor variable in a linear combination.

As indicated in Table 7, from the three sub-test scores of the entrance examination, the variable which has been first entered the regression model is the English language (X_1) since it has the highest correlation with the dependent variable. This sub-test alone improved the accuracy of prediction

by 22.5 percent. The second independent variable entered in the second step is mathematics which added 2.1 percent and the last variable that entered in the regression model is the teaching aptitude which added 1.0 percent to the explained variance in first semester GPA.

In order to examine the statistical significance of the relative contribution of the sub-tests, the squared semi-partial correlation (increment in R^2) is tested. Accordingly, the English language (X_1) contributed 22.5 percent and was found to be highly significant ($F(1,581) = 168.666$ $P < 0.05$). The second sub-test mathematics (X_2) contributed 2.1 percent and was also found to be significant ($F(1,580)=16.154$ $P<0.05$)) and teaching aptitude (X_3)contributed 1.0 percent and was found to be statistically significant too ($F(1,579)=7.782$ $P<0.05$)).

Again, in order to examine the statistical significance of the relative contribution of the three sub-tests, in combination with ESLCE GPA score and high school average score, the squared semi-partial Correlation (increment in R^2) is tested. As it may be observed from Table 8, the English language (X_1) contributed 22.5 percent and was found to be significant ($F(1,581)=168.666$ $P<0.05$)). The second important variable, high school average score (X_5) which contributed 20.4 percent was found to be also significant ($F(1,580)=207.317$ $P<0.05$)). The third important variable, ESLCE GPA (x_4) which contributed 0.5 percent was found to be significant ($F(1,579) = 5.112$ $P< 0.05$)). The last independent variable, Mathematics (x_2) which added 0.2 percent was found to be statistically not significant ($F(1,578)= 2.049$ $p>0.05$)). The sub-test teaching aptitude (X_3) is omitted from the regression model because it does not contribute significantly to the prediction of the criterion measure. Taken as a whole, the independent variables, English language (X_1), high school average score (X_5) and ESLCE GPA (x_4) are the potential predictors.

Table 7. Summary of Stepwise Multiple Regression Analysis of the three sub-test Score on the dependent variable y.

Step Number	Variables entered	Regression Coefficient	Standard error of Regression Coefficient	Regression Weight (β)	Multiple R	Multiple R^2	Increment in R^2	F-ratio
1	x_1	0.166	0.021	0.354	0.474	0.225	0.225	168.666*
2	x_2	0.058	0.015	0.147	0.496	0.246	0.021	16.154*
3	x_3	0.027	0.010	0.119	.506	0.256	0.010	7.782*

* significant at the .05 level.

Multiple Correlation = 0.506 (R^2=0.256)

F – value = 7.782

Constant = 0.756

Standard error of estimate = 0.487

Regression equation

$Y' = 0.756 + 0.166x_1 + 0.058x_2 + 0.027x_3$

Table 8. Summary of Stepwise Multiple Regression Analysis

Step number	Variables entered	Regression Coefficient	Standard error of regression Coefficient	Regression Weight (β)	Multiple R	Multiple R^2	Increment in R^2	F-Ratio
1	x_1	0.086	0.018	0.183	0.474	0.225	0.225	168.666 *
2	x_5	0.102	0.008	0.485	0.655	0.429	0.204	207.317 *
3	x_4	0.013	0.006	0.073	0.658	0.434	0.005	5.112*
4	x_2	0.023	0.014	0.057	0.660	0.436	0.002	2.049

* Significant at .05 level

Multiple Correlation = 0.660 (R^2 = 0.436)

Constant = 0.249

F-value = 2.049

Standard error of estimate = 0.424

Regression equation

$Y' = 0.249 + 0.086x_1 + 0.023x_2 + 0.013x_4 + 0.102x_5$

4.2 Discussion

In an attempt to evaluate the locally constructed entrance examinations that have been employed for selecting candidate teachers for the Arbaminch Teacher Training Institute, several research questions are posed in the present study which are discussed next based on the results found. Each question is presented and discussed here under.

Question 1: Do the items on the entrance examination differentiate candidates' varying levels of ability?

The test analysis report of the three sub-tests of the entrance examination revealed that most of the items of the entrance examination discriminate well. As observed in Table 1,2, and 3 of the test analysis report, 29 items of the English language, 22 items of the mathematics and 19 items of the teaching aptitude entrance examination were found to have discrimination indices 0.20 and higher which could be considered as good items because a rule of thumb be to require discrimination indices to be 0.20 or higher (Ebel, 1966; 1979) The remaining 11 items of English language, 8 items of mathematics, and 11 items of teaching aptitude entrance examination were found to have discrimination indices below 0.20. A few of these items discriminate negatively and a few do not. They were either miskeyed or intrinsically ambiguous. The findings indicate the need to revise or reject such items.

Concerning item difficulty, 4 items in the English language, 5 items in the mathematics and 4 items in the teaching aptitude entrance examination fall in the vicinity of 0.50 difficulty level which could be considered as the optimal difficulty level. The majority of the items in each sub-test were very easy items and only a few items were very difficult (see Appendix D, E, and F). Furthermore, a lot of items in each sub-test were not well-constructed.

Question 2. To what extent are the locally constructed entrance examination reliable and valid?

Reliability coefficients may serve to estimate the precision of the test as a measuring instrument. Reliability estimates based on a single sitting offer evidence as to the precision of the test itself; these include internal consistency estimates, such as those obtained by the use of the split-half and Kuder-Richardson techniques. These techniques were used in the present study to evaluate the precision of sub-tests of the entrance examination.

From the results summarized in Tables 1, 2, and 3, it has been detected that the three sub-tests, English language, mathematics and teaching aptitude had a split-half reliabilities of 0.70, 0.69 and 0.60,respectively. The internal consistency reliabilities of the three sub-tests obtained by the use of K-R-20 were 0.70, 0.66, and 0.63 respectively. The reliability coefficients of the English language and Mathematics entrance examinations were acceptable for screening or preliminary decisions. This is because any testing application requires an internal consistency reliability of around 0.70 or higher (Kline, 1975; Murphy and David Shofer, 1994).

The obtained reliability Coefficient of the Teaching Aptitude entrance examination is some what lower than is desirable for screening purposes or for preliminary decisions. This lack of satisfactory reliability was likely the result of some factors such as the spread of variation in item difficulties. As it is expressed in the literature section of the present study, maximum reliability is achieved when the items fall in the vicinity of 0.50 difficulty level (Ebel and Damrin, 1958; Anastasi, 1976). Furthermore, Myers, (1962) suggested that to have maximum reliability, tests should be homogeneous with respect to item difficulty. Another factor to be considered was the ambiguity of items. An ambiguous event can be interpreted in more than one ways. Individual examinees interpret ambiguous items differently. Such interpretations tend to be random, and hence they increase error variance and decrease reliability. Some of these factors could be considered as the causes for the low reliability

coefficients of the achievement tests:- English language (X_1) and mathematics (X_2).

With regard to validity, the results of the simple correlation analysis revealed that sub-tests of the entrance examination have low predictive validity. This is consistent with previous findings of others (Alvidres and Whitworth, 1981; Meresa, 1994). This lack of predictive validity was likely the result of several factors. First, the lack of specific criteria for constructing achievement and aptitude entrance examination. Second, the lack of sufficient knowledge for constructing aptitude test items. As it may be observed (see Appendix C), it is difficult to consider the items of teaching aptitude as aptitude test items at all. Third, the very high intercorrelations found among the independent variables. Fourth, the lack of satisfactory reliability coefficient of the test, and fifth, the probable lack of correspondence between test and subject contents at the Teacher Training Institutes.

Question 3 . Does the linear combination of the various instruments i.e., the three sub-tests, ESLCE GPA score, and high school average score have a significant contribution in the prediction of academic success of the trainees?

As Neter and Wasserman (1974) pointed out the usual interpretation of a multiple regression equation depends implicitly on the assumption that the independent variables are not strongly correlated with each other. The scholars indicated that when the independent variables are strongly correlated, several difficulties occur in interpreting and evaluating the regression equation. To solve such statistical limitation, different authors (Neter and Wasserman, 1974; Kerlinger; 1973) suggest the use of stepwise multiple regression analysis. With the above limitations in mind, the results obtained in this study warrants the following interpretations.

The result of regression analysis revealed that the linear combination of independent variables contributed 43.6 percent (R = 0.660) to the explained

variance of first semester GPA. The multiple correlation coefficient of the present study is some what similar to the results obtained by other researchers (Scanell, 1960; Willingham, 1964).

From the sub-tests of the entrance examination, English language (x_1) contributes significantly to the prediction of first semester GPA. However, the achievement test, mathematics (X_2), does not contribute significantly to the prediction of the first semester GPA. This is in consistent with previous findings of others (Alvidres and Whitworth, 1981; Meresa, 1994). Furthermore, the locally constructed teaching aptitude entrance examination (X_3), also does not contribute significantly to the prediction of the criterion measure. This result is inconsistent with previous findings (Lakew, 1972; Mekonnen et.al., 1991; Meresa, 1994).

Besides the sub-tests of the entrance examination, High school average score (X_5) which had the highest correlation (r=0.626) with first semester GPA is more significant and important in predicting the candidate teachers ability to perform in the training institute. This is also consistent with the findings of others (Belay, 1990; Meresa, 1994; Kebede, 1991; Passons, 1967, Chissom and Lanier, 1975). Moreover, the ESLCE GPA (X_4) which added 0.5 percent to the explained variance of the criterion measure contribute significantly too.

Taken as a whole, from the five predictor variables, the English language (x_1), ESLCE GPA (x_4), and high school average score (x_5) contribute significantly to the prediction of academic success of the trainees.

Finally, the conclusion and recommendation of the study will be given.

CHAPTER FIVE

5. CONCLUSION AND RECOMMENDATION

5.1 Conclusion

The main purpose of this study was to evaluate the development and validation of an Entrance examination which has been employed for selecting candidate teachers for the Arbaminch Teacher Training Institute.

The Subjects of the study were 583 candidate teachers who were selected from six zone of Southern Nations Nationalities and Peoples State (SNNPS) in 1998/99 academic year. Corrected answer sheets and results of the candidate teachers on three sub-tests were collected from the Education Bureau of SNNPS and ESLCE GPA, high school average scores, and first semester GPA of the candidate teachers were collected from Arbaminch Teacher Training Institute. Attempts have also been made to evaluate and interprete the collected data applying the following statistical tools:

- Item analysis was made to find out difficulty and discrimination indices of each item. Internal consistency reliability estimates were calculated by using item analysis data.
- Correlation coefficients were computed to analyze the degree of relationship between each independent variable and the criterion measure for the total group.
- Multiple regression analysis was made to examine the contribution of the three-sub-tests, the ESLCE GPA, and high school average score together in predicting the criterion measure.

- Stepwise regression model was also used to check if some independent variables improved the accuracy of prediction.

- F and t- tests were utilized to decide whether the observed proportion of variance accounted for by each predictor variable taken alone or selected composite weight are significant for the total group in the study.

Based on the major findings of the study and the review of related literature, some Conclusions were made.

1. Item analysis data for the three sub-tests of the entrance examination indicated that some items from each test could be revised.

2. Estimation of the reliabilities of the three tests with split-half resulted with reliability indices of 0.70, 0.69, and 0.60 and with K-R-20, reliability indices of 0.70, 0.66, and 0.63 resulted for the English language, mathematics, and teaching aptitude tests respectively. From these, it is concluded that the English language and mathematics have acceptable reliability for screening purpose. The teaching aptitude test has relatively low reliability than the other sub-tests of the entrance examination.

3. The simple correlation analysis result revealed that sub-tests of the entrance examination have low predictive validity.

4. The stepwise multiple regression analyses results revealed that English language entrance examination (X_1), high school average score (X_5), and ESLCE GPA score (X_4) are the best and potential predictors of the criterion measure. However, the two sub-tests,

mathematics (X_2) and the teaching aptitude entrance examination (X_3) do not contribute significantly to the prediction of the criterion measure.

5.2 Recommendations

Based on the results presented in Tables 1 through 8 in this study, the following recommendations are made.

1. The acceptable reliability coefficient and significant predictive validity observed by English language entrance examination to the criterion of first semester GPA for the entire sample studied provide evidence for its continued use as admission criterion. That is to say, for screening applicants to join Teacher Training Institutes in SNNPS, it seems reasonable to give more weight to the English language (X_1) than the teaching aptitude entrance examination (X_3).

2. Eventhough, the reliability coefficient of the mathematics entrance examination is dependable for screening purpose and has positive relationship with the criterion measure, its contribution is not significant in combination with other predictor variables. Hence, it is advisable to make some changes concerning the construction of the items of the test.

3. Utmost care needs to be taken concerning the use of the teaching aptitude test as a selection instrument in the future. If there is the need to check for general teaching aptitude in candidates, expert advice should be secured in the making of such tests.

4. Since the high school average score is the better and potential predictor of the criterion measure than the ESLCE GPA, it seems

reasonable to give more weight to high school average score than the ESLCE GPA.

5. Sub-tests of the entrance examination should be prepared by a committee of experts (test specialists).

6. Conduct additional research by including other Teacher Training Institutes so that the results of the study could be generalized to larger population.

REFERENCES

Alvidres, Maria D. and Whitworth, R.H. (1981). The development and validation of an entrance examination in Mexican University. Educational and Psychological Measurement, 41(2), 503-509.

Anastasi, A. (1976). Psychological Testing (4th ed). New York: Macmillan publishing Company, Inc.

_____. (1988) Psychological Testing (6th ed.). New York: Macmillan Publishing Company Inc.

Applegate, J.H. (1987). Teacher candidate selection: An Overview. Journal of Teacher Education, 38 (2), 2-6.

Austin, G.R. and Garber, H. (eds) (1982). The Rise and Fall of National Test Scores. New York ; Academic Press.

Ayele Meshesha and Langmuir, C. (1971). Test results of applications for evening extension: A problem of assessment and selection, "Technical Report", University Testing Centre, HSIU, Addis Ababa.

Belay Teferra (1990). An Evaluation of the Predictive Validity of Admission Criteria for Teacher Training Institutes in Ethiopia. Unpublished Master's Thesis, School of Graduate Studies, AAU.

Bowers, John E. (1968). "Predicting first year success at the University with the Differential Aptitude Tests", Technical Memorandum (1968-1), University Testing Center, HSIU Addis Ababa.

Brown, F.G. (1983). Principles of Educational and Psychological Testing (3rd ed). New York: CBS College Publishing.

Chase, C.I. (1964). Relative length option and response set in multiple choice items. Educational and Psychological measurement. 24, 861-866.

Chauncey, Henry and Dobbin, John E. (1963). Testing: Its Place in Education Today. New York : Harper and Row Publishers.

Chissom, Brad S. and Lanier, D. (1975). "Prediction of first quarter freshman GPA Using SAT scores of high school grades. Educational and Psychological Measurement, 35(2), 461-463.

Cohen, Loius and Manion, (1994). Research Methods in Education (4th ed.) London : Routledge.

Cohen, R.J., Swerdlik, M.E., and Smith, D.C.(1992). Psychological Testing and Assessment: An Introduction to Tests and Measurement. (2nd ed) California: Mayfield Publishing Company.

Coladarci, Theodore (1986). Accuracy of teacher judgements of students responses to standardized test items. Journal of Educational Psychology, 78(2), 141-146.

Coppedge, Floyd L. (1969). "Relation of selected variables from high school records to occupational and college success," The Journal of Educational Research, 63(2), 71-73.

Cronbach, L.J., and Warrington, W.G. (1952). Efficiency of multiple-choice items as a function of spread of item difficulties. Psychometrita, 17, 127-147.

Cureton, E.E. (1957). The upper and lower twenty-seven percent rule. Psychometrika, 22, 293-296.

D'Agostino, R.B., and Cureton, E.E. (1975). The 27 percent rule revisited. Educational and Psychological Measurement, 25, 41-50.

Dalton, S. (1976). "A Decline in the predictive validity of the SAT and Highschool achievement", Educational and Psychological Measurement, 36, (2), 445-448.

Ebel, Robert L. (1966). Obtaining and reporting evidence on content validity. In Chase, C.I. and Ludlow, H.G. (eds.) Reading in Educational and Psychological Measurement. Boston: Houghton Mifflin Company.

Ebel, Robert L. (1979). Essentials of Educational Measurement (3rd ed.). Englewood Cliffs, NJ: Printice-Hall Inc.

Ebel, R.L., and Damirin, D.E. (1958). Tests and examinations. In Harris, Chester W., and Liba, Marie, R. (eds.) Encyclopedia of Educational Research, (3rd ed.) New York : Macmillan Company.

The ESLCE OFFICE (1985). The Ethiopian School Leaving Certificate Examination Handbook. Addis Ababa: AAU.

Engelhart (1975). Suggestions for writing achievement test exercises. In Payne, D.A., and Mcmorris, R.F. (eds.). Educational and Psychological Measurement: Contribution to Theory and Practice, Morristown: General learning Press.

Feldt, L.S. (1963). Note on the use of extreme criterion groups in item discrimination analysis. Psychometrika, 28, 97-104.

Fraenkel, J.R. and Wallen, N.E. (1993). How to Design and Evaluate Research in Education, (2nd ed.). New York: Mc Graw-Hill inc.

Franz, G., Davis, J.A. and Garcia, D. (1958). "Prediction of grades from Preadmission indices in Georgia tax supported colleges, "Educational and Psychological Measurement, 18(4), 841-844.

Gronlund, N.E. (1981). Measurement and Evaluation in Teaching, (4th ed.). New York: Macmillan Publishing Company, Inc.

Guilford, J.P. (1965). Fundamentals of Statistics in Psychology and Education, (4th ed.) New York: Mc Graw-Hill Book Company.

Haberman, M. 1974). Neeed: new guidelines for teacher candidate selection. Journal of Teacher Education, 25, 234-235.

Henrysson, S. (1971). Gathering, analyzing, and using data on test items. In R.L. Thorndike, (ed.), Educational Measurement, (2nd ed.) Washington, D.C. American Council on Education.

Hoge, Robert D., and Butcher, Robert (1984). Analysis of teacher judgements of pupil achievement levels. Journal of Educational Psychology, 76(5), 777-781.

Hopkins, K.D., and Stanley, J.C. (1981). Educational and Psychological Measurement and Evaluation, New Jersey: Prentice-Hall Inc.. Englewood Cliffs.

Jacobson, S. (1963), "Judical review of college admission policies," The Journal of Higher Education, 34 (8), 432 – 436.

Kebede Aderra (1991). The Prediction of the Achievement of students from High School Grade point Average, ESCLE and Selected Aptitude Tests in the Junior College of Commerce. Unpublished Master's Thesis School of Graduate Studies, AAU.

Kerlinger, Fred N. (1973). Foundations of Behavioral Research (2^{nd}.ed.). New York: Holt, Reinehart and Winston Inc.

King, Michael and King, J. (1972) . "Some correlations of University performance in a developing country. The case of Ethiopia, "The Ethiopian Journal of Education, 5 (2), 20-30.

Kline, Paul (1975). Psychometrics and Psychology. London: A subsidiary of Har Court Brace Jovanovich.

Lakew Woldetekle (1972). "The Differential Effectiveness of Amharic and English Language Predictors of College Success in Ethiopia. unpublished Doctoral Dissertation, Syracuse University, USA.

Madsen, Harold S. (1969). "Statistical analysis of the 1968 ESCLE English language examination," "Ethiopian Journal of Education, 3(1), 2-9.

Mann, Sister M.J. (1961). "The prediction of achievement in a liberal arts college, "Educational and Psychological Measurement, 2(2)481-483

Marso, R.N. (1975) Item arrangement and test performance. In Payne, D.A. and Mcmorris R.F. (eds.). Educational and Psychological Measurement: Contribution to Theory and Practice, Morristown: General learning Press.

Mcmorris, R.F., Brown, J.A. and Pruzek, R.M. (1975). Effects of violating item Construction principles. In payne, D.A., and Mcmorris, R.F. (eds.) . Educational and Psychological Measurement: Contribution to Theory and Practice, Morristown: General learning Press.

Mekonnen Yimer, Darge wole, Daniel Desta and Abdinasir Ahmed (1991). The Comparability of the Predictive Power of the Ethiopian Leaving Certificate Examination with Other Measures: Preliminary Observations. Addis Ababa: Institute of Educational Research, AAU. (Memo).

Melaku Asfaw (1975). "The relationship between the ESLCE GPA's of required subjects and first semester freshman GPA's of dismissed students, "The Ethiopian Journal of Education, 12,48-55..

Meresa Abraha (1994). The Predictive validity of the Admission Criteria Used by Ethiopian Electric Light and Power Authority. Unpublished Master's Thesis, School of Graduate Studies, AAU.

Michael, William B. (1965). "Measurement and Prediction in the College admissions process: Some possible directions for future research," Educational and Psychological Measurement, 25(a), 55-72.

Michael, William B. and Jones, Robert A. (1963). "Stability of Predictive validities of high school grades and of Scores on the Scholastic Aptitude Test of the College Entrance Examination Board for liberal arts students", Educational and Psychological Measurement, 23(2), 375-378.

Michael, William B., Jones Robert A., Cox, Anna, Gershon, Arthur, Hoover, Marvin, Katz, Kenneth and Smith, D. (1962). "High School record and College board scores as Predictors of success in a liberal arts Program during the freshman year of College", Educational and Psychological Measurement, 22(2), 399 - 400.

Ministry of Education (1972-1989). Guidelines for teacher candidate selection. Unpublished. (in Amharic).

Mittman, A. (1972). "A multivariate Prediction study for Haile Selassie I University," Technical Report, University Testing Center, HSIU.

Murphy, K.R., and Davidshofer, C.O. (1994). Psychological Testing: Principles and Applications (3rd ed.). New Jersey: Prentice-Hall, Englewood Cliffs.

Myers, C. (1962). The relationship between item difficulty and test validity and reliability, Educational and Psychological Measurement, 22(3), 565-571.

Neter, John and Wasserman, William (1974). Applied Linear Statistical Models. Illionis: Richard D. Irwin, Inc.

Nitko, A.J. (1983). Educational Tests and Measurement: An Introduction, New York: Har Court Brace Jovanovich, Inc.

Passons, William R. (1967). "Predictive validities of the ACT, SAT and high school grades for first semester GPA and freshman courses," Educational and Psychological Measurement, 16(3), 352-400.

Sax, Gilbert (1974). Priniciples of Educational Measurement and Evaluation, California; Wadsworth Publishing Company, inc.

Scanell, Dale, P.(1960). "Prediction of College Success from Elementary and Secondary School Performance, "Journal of Educational Psychology, 51(3), 130-134)

Shenkute Mammo (1991). The Predictive Validity of Selection Criteria of Ethiopian Airlines Aviation Maintenance Technicians School (EAL-AMTS). Unpublished Master's Thesis, School of Graduate Studies, AAU.

Spaulding, H. (1959). "The Prediction of first year grade average in private junior College, "Educational and Psychological Measurement,

Tamiru Geleta (1992). Predictive Validity of the Selection Criteria of Catering and Tourism Training Institute. Unpublished Master's Thesis, School of Graduate Studies, AAU.

Appendix A

Teacher Training Institutes of SNNPS Entrance Examination

Subject- ENGLISH

Name _____ Sex _____ Time Allowed 1Hr.

zone _____ Woreda _____

PART I

GRAMMAR AND USAGE

DIRECTION: DECIDE WHICH ONE OF THE FOLLOWING CHOICES COMPLETES THE SENTENCES BEST.

_____ 1. We must finish now, ____ it's nearly bedtime.

a) by b) for c) on d) at

_____ 2. The rain is getting _____ and _____.

a) heavy/heavier c) heavier/ heavy

b) heavy/ heavy d) heavier/ heavier

_____ 3. _____ the students has scored above fifty.

a) All of b) One of c) Some of d) Many of

_____ 4. Let's ____ to the next topic.

a) move on b) come down c) put away d) go with

_____ 5. She told me that her brother has been living in Jinka _ 1985 E.C.

a) about b) at c) since d) for

_____ 6. _____ some milk in the glass.

a) There are b) Their are c) Their is d) There is

_____ 7. Would you mind _____ that window, please?

a) open b) to open c) opening d) opened

_____ 8. His suggestion is _____ of all his colleagues.

a) the worst b) worse c) bad d) better

Appendix A

Teacher Training Institutes of SNNPS Entrance Examination

Subject- ENGLISH

Name _____ Sex _____ Time Allowed 1Hr.

zone _____ Woreda _____

PART I

GRAMMAR AND USAGE

DIRECTION: DECIDE WHICH ONE OF THE FOLLOWING CHOICES COMPLETES THE SENTENCES BEST.

_____ 1. We must finish now, ____ it's nearly bedtime.

a) by b) for c) on d) at

_____ 2. The rain is getting _____ and _____.

a) heavy/heavier c) heavier/ heavy

b) heavy/ heavy d) heavier/ heavier

_____ 3. _____ the students has scored above fifty.

a) All of b) One of c) Some of d) Many of

_____ 4. Let's ____ to the next topic.

a) move on b) come down c) put away d) go with

_____ 5. She told me that her brother has been living in Jinka _ 1985 E.C.

a) about b) at c) since d) for

_____ 6. _____ some milk in the glass.

a) There are b) Their are c) Their is d) There is

_____ 7. Would you mind _____ that window, please?

a) open b) to open c) opening d) opened

_____ 8. His suggestion is _____ of all his colleagues.

a) the worst b) worse c) bad d) better

_____ 9. We want to _____ peace.

a) do b) making c) make d) doing

_____ 10. Those are my dogs. They belong to _____.

a) I b) mine c) me d) my

DIRECTION: THERE ARE INCOMPLETE CONVERSATIONS BELOW. COMPLETE THE FOLLOWING CONVERSATIONS BY CHOOSING THE BEST ANSWERS FROM THE CHOICES.

_____ 11. WORKU: Hello! kasech. How are you?

KASECH: _____ fine thank you. And you?

a) You're b) We're c) I'm d) He's

_____ 12. WORKU: Fine. I didn't see you last week ____ the bus station. It was ____ the after noon.

a) at/in b) in/in c) at/on d) in/on

_____ 13. KASECH: _____. I was not around that area. I just came from Jimma yesterday.

a) Yes, you have c) No, you didn't

b) Yes, you did d) No, you haven't

_____ 14. WORKU: _____ did you go to Jimma? Can you tell me the reason?

a) Why b) Which c) Where d) When

_____ 15. KASECH: Of course, _____ Jimma has better facilities than Alaba

a) to bring b) to visit c) to converse d) to forget

_____ 16. WORKU: What a nice! I haven't been ____ in Jimma ____ in Alaba.

a) both/ or b) neither/ nor c) nor/but d) either/or

_____ 17. KASECH: If you want, we ___ there together.

a) will go b) went c) had gone d) should have gone

_____ 18. WORKU: _____, but don't forget it.

a) Wrong thing b) Nice idea c) Bad thing d) Which one

_____ 19. KASECH: No I _____. Rather we will arrange a programme for it.

a) was b) do c) won't d) would

_____ 20. WORKU: _____. Try your best.

a) Good of you

b) Impossible

c) Forget it

d) No, thank you

TABLE

NAME	SEX	AGE	WEIGHT	GRADE	BIRTH PLACE
ZUFAN	F	19	48KG	12E	AWASSA
ALEMU	M	18	45KG	10C	SODDO
MEGERSA	M	20	50KG	11B	JINKA
LUBABA	F	17	47KG	12A	DILLA
BOGALE	M	19	51KG	11F	A/MINCH
HAILU	M	1.7	46KG	11D	A/ABABA

DIRECTION: Study the above TABLE very carefully and choose the best answer to complete the sentences below.

_____ 21. The oldest student is from ____ and the youngest girl is from ___

a) Dilla/ Awassa

b) Addis Abeba/Soddo

c) Jinka/ Dilla

d) A/Minch/ Jinka

____ 22. There are more students from grade _____ in the table.

a) 12 b) 10 c) 11F d) 11

____ 23. ____ weighs the least and _____ weighs the most in the table.

a) Alemu/ Bogale

b) Hailu/ Alemu

c) Lubaba/ Zufan

d) Bogale/ Hailu

_____ 24. The student from the lowest grade is from _____

a) Dilla b) Addis Ababa c) Soddo d) Awassa

_____ 25. The number of _____ is greater than that of _____

a) females/ males

b) males/ males

c) females/ females

d) males/ females

____ 26. There are ____ teenagers in the table.

a) 5 b) 4 c) 3 d) 2

_____ 27. The student from Jinka is the highest in _____.

a) age b) weight c) both d) grade

_____ 28. The average grade in the table is _____.

a) 12 b) 11 c) 10 d) 12A

_____ 29._____ and _____ are equal in age.

a) Zufan/ Hailu

b) Bogale/ Alemu

c) Megersa/ Lubaba

d) Lubaba/ Hailu

_____ 30. The oldest student is in grade _____.

a) 11 b) 10 c) 12 d) 12E

PART II: READING COMPREHENSION

TAKING PART IN A DISCUSSION

Come to the discussion prepared with facts, opinions and a discussion out line. Be prepared to cooperate with the rest of the group, to pool your

information, and to work with your class mates to ward the best solution of the problem. Listen carefully. Make sure that each remark you offer relates to the point immediately under discussion. If some one disproves or rejects one of your pet ideas, accept the criticism in good spirit.

Extracted from English In Action.

DIRECTION: Choose the most appropriate answer based on the above passage

_____ 31. Before you join a discussion _____.

a) try to solve the problem yourself
b) disprove your ideas by yourself
c) expect some one may reject your idea
d) don't expect any criticism

____ 32. During a discussion _____

a) take much time for yourself
b) say nothing about the matter
c) listen to others carefully
d) talk unnecessary things

____ 33. When a discussion takes place _____.

a) work with co-workers
b) don't cooperate with others
c) offer irrelevant opinions
d) ignore others

_____ 34. A genuine criticism may _____ your idea in a discussion.

a) damage
b) modify
c) hurt
d) spoil

____ 35. The best solution for any problem may be obtained through ____.

a) force
b) rejection
c) discussion
d) relation

PART 3: SENTENCE COMPREHENSION

DIRECTION: Read the following sentences carefully and then choose the alternative with the same or nearly the same meaning with the original sentence.

_____ 36. Zeritu's body trembled with anger.

a) Zeritu troubled her body unnecessarily
b) Zeritu shaked filling with annoyance.
c) Zeritu traveled a distance carrying her body.
d) Zeritu's body troubled her a lot.

_____ 37. Honey helps to get rid of runny nose.

a) honey controls the nose from running
b) honey helps to have long nose
c) honey can be put in nose
d) honey helps to stop the fluids from nose

_____ 38. Mulu is ahead of Sileshi.

a) Sileshi follows Mulu.
b) Mulu goes after Sileshi.
c) Sileshi and Mulu are equal in speed.
d) Mulu follows sileshi.

_____ 39. Our society is multilingual.

a) Our people have no language.

b) Our people speak many languages.

c) We have a multi-system TV.

d) Multilingual society is lucky.

_____ 40. Almaz couldn't pass ESLCE even though she tried her best.

a) Almaz studied hard and passed ESLCE.

b) Although Almaz had worked hard, she hasn't passed ESLCE.

c) If Almaz worked hard, she would pass ESLCE.

d) Unless Almaz studies hard, she won't pass ESLCE.

Appendix B

Teacher Training Institutes of SNNPS Entrance Examination

Subject- MATHEMATICS

INSTRUCTION: this exam. Contains only 30 multiple choice questions. Each question is followed by five suggested alternatives denoted by the letters a,b,c,d, and e. Choose the best answer from the given alternatives and write the letter of your choice on the space provided on the separate answer sheet.

_____ 1. Which of the following is not a sub set of $\{\emptyset, \{0\}\}$.

a) $\{\emptyset\}$ b) $\{\{0\}\}$ c) $\emptyset$ d) $\{0\}$ e) None of the above

_____ 2. Suppose A is a set with 8 elements, B is a set with 5 elements and the set $A \cap B$ has 3 elements. Then the number of elements in the set $A \cup B$ is

a) 16 c) 10

b) 20 d) 13 e) None of the above

_____ 3. If $A = \{-1,0, \{-11\}\}$ and $B = \{-1,0,1\}$ then which of the following is true?

a) $A \cap B = B$ c) $A \cup B = A$

b) $A \setminus B = \emptyset$ d) $B \setminus A = \emptyset$ e) $A \cap B \setminus \{-1,0\} = \emptyset$

_____ 4. Given any two sets A and B, which of the following is never true?

a) $A \cap B = A$

b) $A \cap B = A$

c) $A \cup B \subseteq A \cap B$

d) $(A \setminus B) \cap (B \setminus A) \neq \emptyset$

e) None of the above

_____ 5. Given A = {∅, {∅},}, B = {{∅, {∅}},∅} C= ∅ which of the following is true?

a) $A \cap B = C$

b) $A \cup B = B$

c) $A \setminus B = C$

d) $B \setminus A = \{\{\emptyset\}\}$

e) $A \Delta B = \{\{\emptyset. \{\emptyset\}\}, \{\emptyset\}\}$

_____ 6. Which of the following are the factors of $X^3 + y^3$

a) $(x-y)(x^2-xy+y^2)$

b) $(x+y)(x^2-xy+y^2)$

c) $(x+y)(x^2+xy+y^2)$

d) $(x-y)(x^2+xy+y^2)$

e) Non of the above

_____ 7. Which of the following is the solution set of the equation $2x - \sqrt{6x-3} \ ?$

a) {1}

b) $\{5\frac{1}{2}\}$

c) $\{1, 5\frac{1}{2}\}$

d) ∅

e) None of the above

_____ 8. If $y = ax^2 + bx + C$, and Y = 2 when x = 0, y = 1 when x = 1, and y = 4 when x = 2, then which of the following is not true?

a) a =2

b) c = 2

c) b = -3

d) $x = \frac{1}{2}$ when y = 0

e) x = 1 when y = 1

_____ 9. If m and n are positive integers, then which of the following is necessarily equal to n^{m^2}?

a) $(n^m)^2$

b) $n^m.n^m$

c) $(n^2)^m$

d) $(n^{m)m}$

e) $n^4\underline{m}2m$
n

_____ 10. A car travels 400 miles on 20 gallons of gasoline. At this rate, how many gallons of gasoline will be consumed on a trip of 900 miles.

a) 35 gallons

b) 55 gallons

c) 45 gallons

d) 65 gallons

e) 25 gallons

_____ 11. The solution set of $\frac{x+1}{x+2} \leq 0$ is:

a) $(\infty, -2) \cup [1,\infty)$

b) $(-\infty, -2]$

c) $[-1, \infty)$

d) $(-2, -1]$

e) None of the above

_____ 12. Which of the following is irrational number?

a) $\sqrt{169}$ c) 0.12112111211112... e) b and c

b) π d) 2.4133333333...

_____ 13. $(120)_{\text{five}} - (101)_{\text{three}}$ is equal to:

a) $(110)_{\text{five}}$

b) $(41)_{\text{five}}$

c) $(11001)_{\text{two}}$

d) $(10101)_{\text{two}}$

e) None of the above

_____ 14. If a is an integer and $a < 0$, then which of the following is true?

a) $|a| = a$

b) $|a| - a = 0$

c) $|a| = -a$

d) $-|a| = -a$

e) None of the above

_____ 15. Which of the following is not true about arithmetic progression whose six and tenth terms respectively 18 and 30.

a) the common difference is 3.

b) the first term is 3.

c) the sum of the first one hundred term is 15,150

d) the sum of any n term is multiple of 3.

e) the arithmetic means of the first ten terms is 30.

_____ 16. If (y323)five = (66y) seven, then y is equal to:

a) 4 b) 0 c) 3 d) 2 e) 1

_____ 17. If $x^3 - y^3 = -2$ and $x^2 + xy + y^2 = 1$, then which of the following is <u>not true</u>?

a) $x^2 = -y^2 + 2xy + 4$

b) $y = x + 2$

c) $-x = 2-y$

d) $x^2 = 2yx + y^2 + 4$

e) None of the above

_____ 18. If $(5^x)^3.8^x = 10\sqrt{10}$, then the value of x is

a) $\frac{1}{4}$ b) $\frac{1}{2}$ c) $\frac{1}{3}$ d) $\frac{1}{6}$ e) None of the above

_____ 19 The GCF and LCM of two numbers are 33 and 330 respectively. If one of the number is 165, then what is the second number?

a) 55 b) 60 c) 66 d) 56 e) 125

_____ 20. Which of the following pair of number is not relatively prime?

a) 9 and 4

b) 10 and 15

c) 7 and 91

d) 8 and 15

e) None of the above

_____ 21. Which of the following is the prime factorization of 4,830?

a) $2^2 \times 3^2 \times 5^2$

b) $2^3 \times 3^2 \times 17$

c) $2 \times 3 \times 5 \times 7 \times 23$

d) $2^2 \times 3 \times 5 \times 7 \times 23$

e) None of the above

_____ 22. Two angles are complementary and the measure of one angle is 20° larger than the measure of the other. How many degrees are there in each angle?

a) 35° and 55°

b) 30° and 60°

c) 35° and 45°

d) 25° and 15°

e) None of the above

_____ 23. In figure 1, area of shaded region is:

(AB = diameter of the circle 0)

a) $(80\pi - 7)cm^2$

b) $(\frac{25\pi}{4} - 6)cm^2$

c) $(\frac{23\pi}{4} - 6)cm^2$

d) $(\frac{2\pi}{4} - 6)cm^2$

e) None of the above

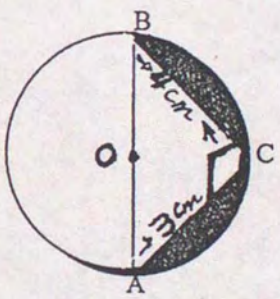

fig. 1

_____ 24. In fig. 2. AB is a straight line. What is the value of x?

a) 25

b) 65

c) 55

d) 45

e) 35

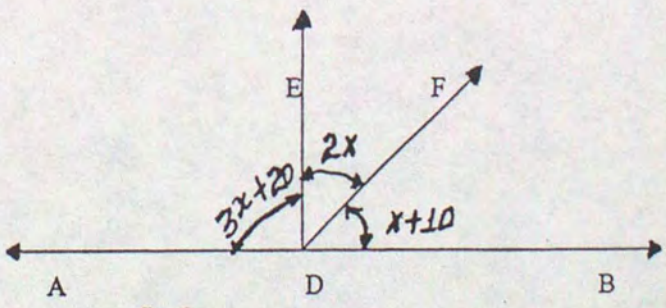

fig. 2

_____ 25. Which of the following triple of numbers can be the length of the sides of a triangle?

a) (1,2,3)

b) (6,2,4)

c) (4,1,5)

d) (3,5,7)

e) None of the above

_____ 26. In fig. 3 if $\overline{AB}$ = 5cm $\overline{AC}$ = 8cm and m (<A) = 60°, then the length of BC: ____?

a) 6cm

b) 7cm

c) 8cm

d) 9cm

e) None of the above

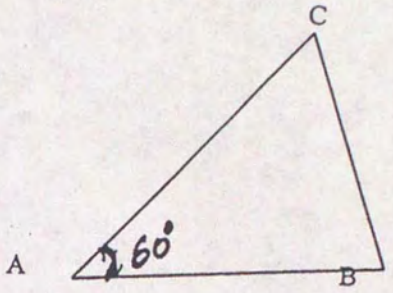

fig. 3

____ 27. The score for five students on a quiz are 40, 20, 30,25 and 15. The mean for the set of score is:

a) 26

b) 30

c) 40

d) 35

e) 50

_____ 28. The probability of obtaining a number greater than 4 on a single toss of a die is:

a) 3

b) $3\frac{1}{2}$

c) $\frac{1}{3}$

d) $\frac{1}{2}$

e) None of the above

_____ 29. Let * and Δ be defined on the set of real numbers by $a * b = a^2 + 2ab$ and $a \Delta b = b/a$, $(a \neq 0)$, then $(c*b) \Delta (a \Delta c)$, $a,b,c \in R$

a) $c^2 + 2cb$

b) $a + 2ab$

c) $a^2 + 2ab$

d) $ac + 2ab$

e) $ac + 2ab$

_____ 30. Find the area of square whose length of diagonal is 8 cm.

a) $36cm^2$

b) $34cm^2$

c) $49cm^2$

d) $32cm^2$

e) None of the above

Appendix C

Teacher Training Institutes of SNNPS Entrance Examination

Subject- TEACHING APTITUDE TEST

Code No. _____

Name _____ Sex _____

zone _____ Woreda _____ Age _____

INSTRUCTION: Read the following sentences very carefully choose the best answer from the given alternatives and write the letter of your choice on the answer sheet.

_____ 1. For which grade level you expect to be a teacher?

a) Grade 1-4
b) Grade 1-6
c) Grade 1-8
d) Grade 5-8

_____ 2. Assume that you are a teacher in a primary grade. As you enter the class, you see an ugly picture drawn on the chalk board with your name at the top. What is your first measure that you take?

a) I punish all of the students in the class.
b) I ask who draw the picture, then I punish him.
c) Understanding the reason why the picture is drawn, and to take a corrective measure.
d) Students are not expected to draw such type of a picture.

_____ 3. As a teacher of a primary school, your most relationship is with

a) the director
b) parents
c) students
d) other teachers

_____ 4. The most common type of behaviour among primary school children.

a) disagreement with the teachers.
b) giving more emphasis for playing
c) leaving the class while the teacher is teaching
d) interest in opposite sex

_____ 5. Assume that you are a teacher in one of the rural schools. What will be your measure if you are invited to eat a food that you dislike or not know before

a) I will advise them not to eat such type of food.
b) I will eat the food for the interest of others.
c) I will ask excuse that I will not eat because of personal problems.
d) No one is concerned whether I eat the food or not.

_____ 6. What is your first measure if your salary is not paid in time?

a) I stop my work, and send the children to their home.
b) I do not worry whether my salary is paid or not.
c) I collect some money from my students.
d) continuing my work, I ask responsible bodies to get my salary.

_____ 7. One of the following is not a desirable behaviour of a good teacher

a) Knowing the name of his students
b) allowing students to discuss on academic problems
c) punishing students who give answer incorrectly
d) rewarding students for their correct answer

_____

b) I fight with him since he has no any right to call me "Mr. Chalk".
c) I discuss with him calling by such name has a negative impact on the profession.
d) I try to give him a name which is related to his profession.

_____ 9. When do you leave your profession?

a) When I get chance for further learning.

b) When I get a better job opportunity.

c) If I have no chance to be transferred to urban areas.

d) If I am assigned to teach in a very remote area.

_____ 10. Which statement is correct?

a) Females are better in teaching music than males.

b) Males are better in teaching physical education than females.

c) Effectiveness of teaching in both subjects depends on the quality of the teachers.

d) Females are not effective in teaching both music and physical education.

_____ 11. On which idea do you agree?

a) Beautiful girls are more clever than others.

b) Beautiful boys are more clever than others.

c) Ugly boys and girls are more clever than others.

d) To be beautiful or ugly is not a cause for cleverness or weakness.

_____ 12. When do you think that children hate their teacher? If he,

a) has poor personality

b) is punctual

c) knows what he teaches

d) respects his students

_____ 13. If you are assigned to teach a subject that you do not like, what is your response?

a) I ask the director to change the subject

b) It is my task to teach all subjects

c) The director has no right to assign me to teach a subject that I do not like.

d) I ask the help of other teachers to teach the subject that I do not like.

_____ 14. Which one is the supply that you do not expect from the Teacher Training Institute?

a) Money
b) Food
c) Dormitory
d) Medical service

_____ 15 Assume that you are a natural science student in your secondary high school. However, you learn social science subjects in the T.T.I. How do you solve this problem?

a) I do not accept to learn subjects out of my field.
b) I leave the class when the social science teacher comes.
c) I agitate students not to learn social science subjects.
d) I agree to learn both the natural science and social science subjects.

_____ 16. "Teaching is the noblest profession" why? because,

a) All teachers are graduates of Teacher Training Institute.
b) Result of teaching profession is the only means for further development.
c) All teachers are successful in their profession.
d) Teaching is the means of getting employment and salary.

_____ 17. As a teacher of grade one, what is your first task as you enter to the class?.

a) Correcting the assignment given.
b) To say "good morning" or "good after noon".
c) To send home, students who were absent.
d) To say "good buy, good bye".

_____ 18. If you are tired and unable to enter the class, what do you do?

a) What ever it is, I never miss the class.
b) I allow students to return to their home.
c) I ask the director for permission.

d) I organize the class to discuss for them selves on an educational issue

____ 19. Which one of the following is arranged in a sequential order to teach social science in primary grades.

a) Family, Woreda, Kebele, Zone
b) Zone, Woreda, Family, Kebele
c) Family, Kebele, Woreda, Zone
d) Woreda, Kebele, zone, Family

____ 20. One of the following material is not used commonly for all subjects in the primary grades.

a) text book
b) chalk-board
c) pen
d) radio

____ 21. Regarding misbehaviour, one of the following affects the teaching learning process as well as the relationship among students.

a) failure to show interest
b) sleeping in the class
c) inattentivness
d) shouting in the class

____ 22. Which factor results class room misbehaviour that is caused by the teacher.

a) poor teaching ability c) large class size
b) lack of interest in learning d) students health problem

____ 23. The activity given for the students should be

a) to keep the class silent
b) to make students creative
c) to make students busy
d) to punish students

d) I organize the class to discuss for them selves on an educational issue

_____ 19. Which one of the following is arranged in a sequential order to teach social science in primary grades.

a) Family, Woreda, Kebele, Zone
b) Zone, Woreda, Family, Kebele
c) Family, Kebele, Woreda, Zone
d) Woreda, Kebele, zone, Family

_____ 20. One of the following material is not used commonly for all subjects in the primary grades.

a) text book
b) chalk-board
c) pen
d) radio

_____ 21. Regarding misbehaviour, one of the following affects the teaching learning process as well as the relationship among students.

a) failure to show interest
b) sleeping in the class
c) inattentivness
d) shouting in the class

_____ 22. Which factor results class room misbehaviour that is caused by the teacher.

a) poor teaching ability
b) lack of interest in learning
c) large class size
d) students health problem

____ 23. The activity given for the students should be

a) to keep the class silent
b) to make students creative
c) to make students busy
d) to punish students

____ 24. A good teacher is one who

a) takes too much time in talking

b) makes friendly comments

c) shows sour facial expression

d) is strict and not flexible

_____ 25. Teachers have to direct their questions

a) to students who sit in front of the class

b) to students who sit at the back

c) to students who sit in the line of their sight

d) to all students equally

_____ 26. If there are no teaching materials, a teacher should

a) ask another teacher with more experience to teach the lesson

b) create materials from resources available in the area

c) stop teaching until the materials are supplied by the school

d) teach the lesson with out the materials because the lesson is important

____ 27. If the student always disturbs the class, what is the best way to deal with that student?

a) punishing the student using different materials until the correct behaviour is obtained

b) giving advice and if that fails, to tell the student not to come back to the class.

c) giving advice and if that fails to contact with parents to discuss the problems

d) asking the school director to transfer the student to another school

_____ 28. What will you do if more than one student give an answer at the same time to your question? I will,

a) punish the students not to give an answer in the future

b) advise them to give an answer when they are asked

c) accept heir answer

d) not give attention to the students

_____ 29. One of the following is not good for teacher-student relationship.

a) ability to share feeling and opinion of the students

b) suppressing the idea of the students

c) not to reject the answer, even if it is a wrong answer

d) all are good for teacher-student relationship

_____ 30. Which one of the following is the way to respond to the parent who complains about the grade given to his/ her child?

a) I do not listen parents for their complain

b) I punish the student for involving his/ her parent

c) I explain to parents how the grade is given

d) it is not the right of parents to ask about students grade.

Appendix - D

Item Discrimination, Difficulty, and Reliability Indices for English Language Entrance Examination.

Item Number		Number of correct Responses for groups (n=157) n=27% x 583 = 157	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
1	H	60	$P_H = 0.38$	D=0.13	P=0.32
	L	40	$P_L = 0.25$		
2	H	142	$P_H=0.90$	D=0.62	P= 0.59
	L	44	$P_L = 0.28$		
3	H	135	$P_H=0.86$	D=0.52	P=0.60
	L	54	$P_L=34$		
4	H	102	$P_H=0.65$	D=0.31	P=0.50
	L	54	$P_L=0.34$		
5	H	154	$P_H=0.98$	D=0.35	P=0.81
	L	99	$P_L=0.63$		
6	H	141	$P_H=0.90$	D=0.42	P=0.69
	L	75	$P_L=0.48$		
7	H	112	$P_H=0.71$	D=.50	P=0.46
	L	33	$P_L=0.21$		
8	H	152	$P_H=0.97$	D=0.47	P=0.74
	L	78	$P_L=0.50$		

Item Number		Number of correct Responses for groups (n=157) n=27% x 583 = 157	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
9	H	119	P_H=0.76	D=0.31	P=0.61
	L	70	P_L=0.45		
10	H	125	P_H=0.80	D=0.33	P=0.64
	L	74	P_L=0.47		
11	H	156	P_H=0.99	D=0.11	P=0.94
	L	138	P_L=0.88		
12	H	110	P_H=0.70	D=0.38	P=0.51
	L	51	P_L=0.32		
13	H	22	P_H=0.14	D=0.00	P=0.14
	L	22	P_L=0.14		
14	H	155	P_H=0.99	D=0.19	P=0.90
	L	125	P_L=0.80		
15	H	145	P_H=0.92	D=0.29	P=0.78
	L	99	H_L =0.63		
16	H	53	P_H=0.34	D=0.07	P=0.32
	L	43	P_L=0.27		
17	H	151	P_H=0.96	D=0.39	P=0.77
	L	90	P_L=0.57		

Item Number		Number of correct Responses for groups (n=157) n=27% x 583 = 157	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
18	H	153	P_H=0.97	D=0.38	P=0.78
	L	93	P_L=0.59		
19	H	94	P_H=0.60	D=0.34	P=0.43
	L	41	P_L=0.26		
20	H	124	P_H=0.79	D=0.36	P=0.61
	L	67	P_L=0.43		
21	H	153	P_H=0.97	D=0.21	P=0.87
	L	120	P_L=0.76		
22	H	138	P_H=0.88	D=0.37	P=0.70
	L	80	P_L=0.51		
23	H	152	P_H=0.97	D=0.26	P=0.84
	L	112	P_L=0.71		
24	H	155	P_H=0.99	D=0.16	P=0.91
	L	130	P_L=0.83		
25	H	156	P_H=0.99	D=0.17	P=0.91
	L	128	P_L=0.82		
26	H	61	P_H=0.39	D=0.10	P=0.34
	L	45	P_L=0.29		

Item Number		Number of correct Responses for groups (n=157) n=27% x 583 = 157	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
27	H	150	PH=0.96	D=0.23	P=0.85
	L	115	P_L=0.73		
28	H	124	PH=0.79	D=0.23	P=0.68
	L	88	P_L=0.56		
29	H	155	P_H=0.99	D=0.11	P=0.94
	L	138	P_L=0.88		
30	H	154	P_H=0.98	D=0.27	P=0.85
	L	112	P_L=0.71		
31	H	64	P_H=0.41	D=0.10	P=0.36
	L	49	P_L=0.31		
32	H	153	P_H=0.97	D=0.24	P=0.85
	L	114	P_L=0.73		
33	H	124	P_H=0.79	D=0.36	P=0.61
	L	67	P_L=0.43		
34	H	111	P_H=0.71	D=0.07	P=0.68
	L	100	P_L=0.64		
35	H	153	P_H=0.97	D=0.21	P=0.87
	L	120	$P_{L=}$0.76		

Item Number		Number of correct Responses for groups (n=157) n=27% x 583 = 157	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
36	H	100	P_H=0.64	D=0.36	P=0.46
	L	44	P_L=0.28		
37	H	134	P_H=0.85	D=0.23	P=0.74
	L	97	P_L=0.62		
38	H	145	P_H=0.92	D=0.44	P=0.70
	L	76	P_L=0.48		
39	H	151	P_H=0.96	D=0.30	P=0.81
	L	103	P_L=0.66		
40	H	130	P_H=0.83	D=0.47	P=0.60
	L	57	P_L=0.36		

* H= Higher Group D = Discrimination index N = 583

L= Lower Group P = Difficulty index n = 157

Appendix D Continued ....

Calculation of split – half reliability for English Language Entrance Examination.

The summary data for English Language Entrance Exam.

	Σx	Σx^2	$\bar{x}$	s	
Odds ($_o$)	8012	113748	13.742	2.499	$\Sigma x_o x_e$ = 108987
Evens ($_e$)	7768	108214	13.324	2.843	N = 583
Total	15780	221962	27. 07	4.693	

$$r_{xx} = r_{oe} = \frac{\Sigma x_o x_e / N - (\bar{x}_o)(\bar{x}\, e)}{(s_o)\ (s_e)}$$

$$= \frac{10897/583 - (13.742)\ (13.324)}{2.499 \times 2.843}$$

$$= 0.540951232$$

$$r_{xx'} = \frac{2\, r_{oe}}{1+r_{oe}} = \frac{1.081902465}{1.540951232}$$

$$= 0.702100393$$

$$r_{xx'} = 0.70$$

Appendix D Continued ....

Computation of a coefficient of Internal consistency for English Language Entrance Examination. (n = 583)

Item	Number correct	P	Pq
1	171	0.293	0.207
2	381	0.654	0.227
3	351	0.602	0.240
4	283	0.485	0.250
5	492	0.844	0.132
6	408	0.700	0.210
7	255	0.437	0.246
8	447	0.767	0.179
9	348	0.597	0.241
10	376	0.645	0.229
11	553	0.949	0.049
12	318	0.545	0.248
13	90	0.154	0.130
14	520	0.892	0.096
15	470	0.806	0.156
16	190	0.326	0.220
17	454	0.779	0.172
18	482	0.827	0.143
19	230	0.395	0.239
20	349	0.599	0.240
21	522	0.895	0.094
22	411	0.705	0.208
23	513	0.880	0.106
24	543	0.931	0.064
25	531	0.911	0.081
26	193	0.331	0.221
27	490	0.840	0.134
28	397	0.681	0.217
29	549	0.942	0.055
30	518	0.889	0.099
31	195	0.334	0.222
32	510	0.875	0.110
33	364	0.624	0.234

34	398	0.683	0.217
35	520	0.892	0.096
36	266	0.456	0248
37	438	0.751	0.187
38	429	0.736	0.194
39	482	0.827	0.143
40	343	0.588	0.242
Total	15780	$\Sigma p = 27.067$ $\Sigma p = 27.07$	$\Sigma pq = 7.026$

$$k\text{-}R\text{-}20 = \frac{k}{k-1}\left(1 - \frac{\Sigma Pq}{S_x^2}\right)$$

$$= \frac{40}{39}\left(1 - \frac{7.026}{21.99}\right)$$

$$\underline{k\text{-}R\text{-}20 = 0.70}$$

$$k\text{-}R\text{-}21 = \frac{k}{k-1}\left(1 - \frac{\bar{x}(k - \bar{x})}{ks_x^2}\right)$$

$$= \frac{40}{39}\left(\frac{1-27.07\,(40-27.07)}{40 \times 21.99}\right)$$

$$= 0.62$$

Appendix - E

Item discrimination, Difficulty, and Reliability Indices for 30 Mathematics Entrance Examination

Item Number		Number of correct Responses for groups (n=157)	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
1	H	136	$P_H = 0.87$	D=0.59	P=0.57
	L	44	$P_L = 0.28$		
2	H	155	$P_H=0.99$	D=0.51	P=0.74
	L	76	PL=0.48		
3	H	139	$P_H=0.89$	D=0.55	P=0.62
	L	54	$P_L=0.34$		
4	H	69	$P_H=0.44$	D=0.21	P=0.34
	L	36	$P_L=0.23$		
5	H	51	PH=0.32	D=0.10	P=0.27
	L	35	$P_H=0.22$		
6	H	127	$P_L=0.81$	D=0.57	P=0.53
	L	38	$P_L=0.24$		
7	H	33	$P_H=0.21$	D=0.06	P=0.18
	L	23	$P_L=0.15$		
8	H	94	$P_H=0.60$	D=0.27	P=0.47

Item Number		Number of correct Responses for groups (n=157)	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
	L	52	P_L=0.33		
9	H	146	P_H=0.93	D=0. 66	P=0.60
	L	43	P_L=0.27		
10	H	145	P_H=0.92	D=0.24	P=0.80
	L	106	P_L=0.68		
11	H	81	P_H=0.52	D=0.20	P=0.42
	L	50	P_L=0.32		
12	H	83	P_H=0.53	D=0.17	P=0.45
	L	57	P_L=0.36		
13	H	140	P_H=0.89	D=0.59	P=0.60
	L	47	P_L=0.30		
14	H	36	P_H=0.23	D=-0.04	P=0.25
	L	42	P_L=0.27		
15	H	87	P_H=0.55	D=0.38	P=0.36
	L	27	P_L=0.17		
16	H	138	P_H=0.88	D=0.47	P=0.65
	L	65	P_L=0.41		
17	H	34	P_H=0.22	D=0.09	P=0.18
	L	20	P_L=0.13		

Item Number		Number of correct Responses for groups (n=157)	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
18	H	99	P_H=0.63	D=0.40	P=0.43
	L	36	P_L=0.23		
19	H	112	P_H=0.71	D=0.37	P=0.53
	L	53	P_L=0.34		
20	H	85	P_H=0.54	D=0.06	P=0.51
	L	75	P_L=0.48		
21	H	133	P_H=0.85	D=0.33	P=0.69
	L	82	P_L=0.52		
22	H	109	P_H=0.69	D=0.52	P=0.43
	L	27	P_L=0.17		
23	H	79	P_H=0.50	D=0.38	P=0.31
	L	19	P_L=0.12		
24	H	148	P_H=0.94	D=0.60	P=0.64
	L	54	P_L=0.34		
25	H	120	P_H=0.76	D=0.50	P=0.51
	L	41	P_L=0.26		
26	H	54	P_H=0.34	D=0.23	P=0.23
	L	17	P_L=0.11		

Item Number		Number of correct Responses for groups (n=157)	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
27	H	150	P_H=0.96	D=0.54	P=0.69
	L	66	P_L=0.42		
28	H	80	P_H=0.51	D=0.36	P=0.33
	L	24	P_L=0.15		
29	H	156	P_H=0.99	D=-0.01	P=1.00
	L	157	P_L=1.00		
30	H	126	P_H=0.80		
	L	82	P_L=0.52	D=0.28	P=0.66

H = Higher group

L = Lower group

P_H = Proportion of Correct responses of Higher group.

P_L = Proportion of Correct responses of lower group.

D = Discrimination.

P = Difficulty.

Appendix E continued ....

Calculation of split – half reliability for Mathematics Entrance Examination.

The summary data for English Language Entrance Exam.

	Σx	Σx^2	$\bar{x}$	s	
Odds ($_o$)	4453	37397	7.638	2.409	$\Sigma x_o x_e$ = 35326
Evens ($_e$)	4401	36329	7.549	2.308	N = 583
Total	8854	73726	15.187	4.123	

$$r_{xx} = r_{oe} = \frac{\Sigma x_o x_e / N - (\bar{x}_o)(\bar{x}\, e)}{(s_o)\ (s_e)}$$

$$= \frac{35326/583 - (7.638)\ (7.549)}{2.409 \times 2.308}$$

$$r_{oe} = \frac{2.93421999}{5.559972}$$

$$= 0.527740066$$

$$r_{xx'} = \frac{2\, r_{oe}}{1+r_{oe}} = \frac{1.055480132}{1.527740066}$$

$$r_{xx'} = 0.69$$

Appendix E Continued ....

Computation of a coefficient of Internal consistency for Mathematics Entrance Examination. (n = 583)

Item	Number correct	P	Pq
1	351	0.602	0.240
2	455	0.780	0.172
3	376	0.645	0.229
4	182	0.312	0.215
5	137	0.235	0.180
6	328	0.563	0.246
7	103	0.177	0.146
8	268	0.460	0.248
9	370	0.635	0.232
10	480	0.823	0.146
11	248	0.425	0.244
12	255	0.437	0.246
13	360	0.617	0.236
14	153	0.262	0.193
15	181	0.310	214
16	391	0.671	0.221
17	87	0.149	0.127
18	221	0.379	0.235
19	310	0.532	0.249
20	303	0.520	0.250
21	429	0.736	0.194
22	247	0.424	0.244
23	175	0.300	0.210
24	400	0.686	0.215
25	297	0.509	0.250
26	129	0.221	0.172
27	451	0.774	0.175
28	182	0.312	0.215
29	580	0.995	0.005
30	405	0.695	0.212
	8854	$\Sigma p = 15.187$	$\Sigma pq=6.161$

$$k\text{-}R\text{-}20 = \frac{k}{k-1}\left(1 - \frac{\Sigma Pq}{S_x^2}\right)$$

$$= \frac{30}{29}\left(1 - \frac{6.161}{17.003}\right)$$

$$= 0.66$$

$$k\text{-}R\text{-}21 = \frac{k}{k-1}\left(1 - \frac{\bar{x}(k - \bar{x})}{ks_x^2}\right)$$

$$= \frac{30}{29}\left(\frac{1\text{-}15.187\ (30\text{-}15.1874)}{30x17.003}\right)$$

$$=0.58$$

Appendix - F

Item discrimination, difficulty, and Reliability indices for Teaching Aptitude Entrance Examination

Item Number		Number of correct Responses for groups (n=157)	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
1	H L	110 67	P_H = 110/157=0.70 P_L= 67/157 = 0.43	D= 0.27	$\frac{.70 + .43}{2}$ = 0.57
2	H L	153 124	P_H=0.97 PL=0.79	D=0.18	P=0.88
3	H L	154 114	P_H=0.98 P_L=0.73	D=0.25	P=0.86
4	H L	152 91	P_H=0.97 P_L=0.58	D=0.39	P=0.86
5	H L	92 73	PH=0.59 P_H=0.46	D=0.13	P=0.53
6	H L	154 126	P_L=0.98 P_L=0.80	D=0.18	P=0.89
7	H L	150 98	P_H=0.96 P_L=0.62	D=0.34	P=0.79

Item Number		Number of correct Responses for groups (n=157)	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
8	H	157	P_H=1.00	D=0.01	P=1.00
	L	156	P_L=0.99		
9	H	136	P_H=0.87	D=0. 51	P=0.62
	L	56	P_L=0.36		
10	H	152	P_H=0.97	D=0.19	P=0.88
	L	123	P_L=0.78		
11	H	156	P_H=0.99	D=0.12	P=0.93
	L	136	P_L=0.87		
12	H	82	P_H=0.52	D=0.32	P=0.36
	L	32	P_L=0.20		
13	H	112	P_H=0.71	D=0.37	P=0.53
	L	54	P_L=0.34		
14	H	137	P_H=0.87	D=0.51	P=0.62
	L	56	P_L=0.36		
15	H	154	P_H=0.98	D=0.20	P=0.88
	L	122	P_L=0.78		
16	H	116	P_H=0.74	D=0.32	P=0.58
	L	66	P_L=0.42		

Item Number		Number of correct Responses for groups (n=157)	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
17	H	153	P_H=0.97	D=0.06	P=0.94
	L	143	P_L=0.91		
18	H	121	P_H=0.77	D=0.18	P=0.62
	L	74	P_L=0.47		
19	H	155	P_H=0.99	D=0.18	P=0.90
	L	127	P_L=0.81		
20	H	15	P_H=0.10	D=-0.01	P=0.11
	L	17	P_L=0.11		
21	H	67	P_H=0.43	D=0.32	P=0.27
	L	18	P_L=0.11		
22	H	123	P_H=0.78	D=0.44	P=0.56
	L	53	P_L=0.34		
23	H	98	P_H=0.62	D=0.38	P=0.43
	L	38	P_L=0.24		
24	H	110	P_H=0.70	D=0.50	P=0.45
	L	32	P_L=0.20		
25	H	151	P_H=0.96	D=0.29	P=0.82
	L	105	P_L=0.67		

Item Number		Number of correct Responses for groups (n=157)	Proportion of Correct Responses	Item Discrimination $D = P_H - P_L$	Item Difficulty $P = \frac{P_H + P_L}{2}$
26	H	126	P_H=0.80	D=0.38	P=0.61
	L	66	P_L=0.42		
27	H	153	P_H=0.97	D=0.28	P=0.83
	L	108	P_L=0.69		
28	H	155	P_H=0.99	D=0.35	P=0.83
	L	101	P_L=0.64		
29	H	40	P_H=0.25	D=0.10	P=0.20
	L	24	P_L=0.15		
30	H	152	P_H=0.97	D=0.29	P=0.83
	L	106	P_L=0.68		

H = Higher group P_H = Proportion of correct responses of higher group D = Discrimination

L = Lower group P_L = Proportion of correct responses of lower group P = Difficulty

Appendix F continued .....

Calculation of split – half reliability for Teaching Aptitude Entrance Examination.

The summary data for English Language Entrance Exam.

	Σx	Σx^2	$\bar{x}$	s	
Odds ($_o$)	5961	63203	10.225	1.966	$\Sigma x_o x_e = 61930$
Evens ($_e$)	5955	3509	10.214	2.145	N = 583
Total	11916	126712	20.439	3.470	

$$r_{xx} = r_{oe} = \frac{\Sigma x_o x_e / N - (\bar{x}_o)(\bar{x}_e)}{(s_o)(s_e)}$$

$$= \frac{61930/583 - (10.225)(10.214))}{1.966 \times 2.145}$$

$$r_{oe} = 0.424053927$$

$$r_{xx'} = \frac{2\, r_{oe}}{1+r_{oe}} = \frac{2.(0.424053927)}{1.424053927}$$

$$= 0.595558804$$

$$= \underline{0.60}$$

Appendix F Continued ....

Computation of a coefficient of Internal consistency for Teaching Aptitude Entrance Examination. (n = 583)

Item	Number correct	P	Pq
1	332	0.569	0.245
2	517	0.887	0.100
3	508	0.871	0.112
4	487	0.835	0.137
5	287	0.492	0.250
6	538	0.923	0.071
7	482	0.827	0.143
8	582	0.998	0.002
9	359	0.616	0.237
10	517	0.887	0.100
11	555	0.952	0.046
12	203	0.348	0.227
13	318	0.545	0.248
14	381	0.654	0.227
15	522	0.895	0.094
16	345	0.592	0.242
17	555	0.952	0.046
18	380	0.652	0.227
19	543	0.931	0.064
20	49	0.084	0.077
21	151	0.259	0.192
22	326	0.559	0.246
23	257	0.441	0.247
24	238	0.408	0.241
25	495	0.849	0.128
26	379	0.650	0.227
27	495	0.849	0.128
28	496	0.851	0.127
29	109	0.187	0.152
30	510	0.875	0.110
Total	11916	20.438 $\Sigma p=20.44$	4.693=Σpq

$$\text{k-R-20} = \frac{k}{k-1}\left(1 - \frac{\Sigma Pq}{S_x^2}\right)$$

$$= \frac{30}{29}\left(1 - \frac{4.693}{12.040}\right) \qquad s_x^2 = 12.040$$

$$= \underline{0.63}$$

$$\text{k-R-21} = \frac{k}{k-1}\left(1 - \frac{\bar{x}(k - \bar{x})}{ks_x^2}\right)$$

$$= \frac{30}{29}\left(\frac{1-20.44\,(30-20.44)}{30 \times 12.040}\right)$$

$$= 0.4748$$

$$\text{k-R-20} = \underline{0.47}$$

Appendix G

TEST ANALYSIS REPORT FORMAT

Test Titles : _____ Job Number _____

Group Tested: _____ Date of Test _____

Time Limit: _____ Checker _____

Characteristic

I. Relevance

- A. Content details _____ _____
- B. Vocabulary _____ _____
- C. Facts _____ _____
- D. Generalizations _____ _____
- E. Understandings _____ _____
- F. Applications _____ _____

II. Discrimination

A. Item

1. High (.41 and up) _____
2. Moderate (.21 to .40) _____
3. Low (.01 to .20) _____
4. Zero or Negative _____

B. Score

1. Mean _____
2. Standard deviation _____
3. Reliability _____
4. Probable error _____

III. Speeded-ness

A. Percent of complete papers.

(a) Midpoint of range between highest possible and expected chance score.

(b) One-sixth of range between possible and expected chance error.