A STUDY OF THE PREDICTIVE VALIDITY OF THE BAŞKENT UNIVERSITY ENGLISH PROFICIENCY EXAM THROUGH THE USE OF THE TWO-PARAMETER IRT MODEL'S ABILITY ESTIMATES

TANER YAPAR

JUNE 2003

A STUDY OF THE PREDICTIVE VALIDITY OF THE BAŞKENT UNIVERSITY ENGLISH PROFICIENCY EXAM THROUGH THE USE OF THE TWO-PARAMETER IRT MODEL'S ABILITY ESTIMATES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

TANER YAPAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF EDUCATIONAL SCIENCES

JUNE 2003

Approval of the School of Social Sciences

Prof. Dr. Bahattin Akşit
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Hasan Şimşek
Head of the Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Giray Berberoğlu
Supervisor

Examining Committee Members

Prof. Dr. Giray Berberoğlu

Assist. Prof. Dr. Ayşegül Daloğlu

Assist. Prof. Dr. Zeynep Sümer

ABSTRACT

A STUDY OF THE PREDICTIVE VALIDITY OF THE BAŞKENT UNIVERSITY ENGLISH PROFICIENCY EXAM THROUGH THE USE OF THE TWO-PARAMETER IRT MODEL'S ABILITY ESTIMATES

Yapar, Taner

M.S., Department of Educational Sciences

Supervisor: Prof. Dr. Giray Berberoğlu

June 2003, 117 pages

The purpose of this study is to analyze the predictive power of the ability estimates obtained through the two-parameter IRT model on the English Proficiency Exam administered at Başkent University in September 2001 (BUSPE 2001). As prerequisite analyses the fit of one- and two-parameter models of IRT were investigated.

The data used for this study were the test data of all students (727) who took BUSPE 2001 and the departmental English course grades of the passing students.

At the first stage, whether the assumptions of IRT were met was investigated. Next, the observed and theoretical distribution of the test data was reviewed by using chi square statistics. After that, the invariance of ability estimates across different sets of items and invariance of item parameters across different groups of students were examined.

At the second stage, the predictive validity of BUSPE 2001 and its subtests was analyzed by using both classical test scores and ability estimates of the better fitting IRT model.

The findings revealed that the test met the assumptions of unidimensionality, local independence and nonspeededness, the assumptions of equal discrimination indices was not met. Whether the assumption of minimal guessing was met remained vague. The chi square statistics indicated that only the two parameter model fitted the test data. The ability estimates were found to be invariant across different item sets and the item parameters were found to be invariant across different groups of students.

The IRT estimated predictive validity outweighed the predictive validity calculated through classical total scores both for the whole test and its subtests. The reading subtest was the best predictor of future performance in departmental English courses among all subtests.

Keywords: Item Response Theory, Predictive Validity, Ability Estimate, BUSPE 2001, DEC 1-2 scores, Item Information Index, Test Information Curve.

ÖZ

İKİ PARAMETRELİ MADDE TEPKİ KURAMI (MTK) MODELİNİN YETENEK KESTİRİMLERİYLE BAŞKENT ÜNİVERSİTESİ İNGİLİZCE YETERLİK SINAVININ YORDAMA GEÇERLİĞİNİ İNCELEME ÇALIŞMASI

Yapar, Taner

Yüksek Lisans, Eğitim Bilimleri Bölümü

Tez Yöneticisi: Prof. Dr. Giray Berberoğlu

Haziran 2003, 117 sayfa

Bu çalışmanın amacı Başkent Üniversitesinde Eylül 2001'de uygulanmış olan İngilizce yeterlik sınavı (BÜEYS 2001) için iki parametreli MTK modelinde elde edilen yetenek kestirimlerinin yordama kuvvetini analiz etmektir. Önkoşul analiz olarak da MTK'nın bir ve iki parametreli MTK modellerine uygunluğu incelenmiştir.

Bu çalışmada uygunluk analizi için BÜEYS 2001'e giren tüm öğrencilerin (727) oluşturduğu sınav verisi ve başarılı olanların bölüm İngilizcesi notları kullanılmıştır.

İlk aşamada, MTK sayıltılarının karşılanıp karşılanmadığı incelenmiştir. Daha sonra, sınav verisinin gözlenen ve kuramsal dağılımı khi-kare istatistiği kullanılarak gözden geçirilmiştir. Bundan sonra, yetenek kestirimlerinin değişmezliği farklı madde grupları ve madde parametrelerinin değişmezliği ise farklı öğrenci grupları kullanılarak incelenmiştir.

İkinci aşamada BÜEYS 2001in ve alt testlerinin yordama geçerliği hem klasik puanlar hem de iki parametreli MTK modelinin yetenek kestirimleri kullanılarak araştırılmıştır.

Bulgular sınavın tek boyutluluk, maddelerin bağımsızlığı, hız testi olmama sayıltılarını karşıladığını, eşit ayırıcılık indeksleri sayıltısını karşılamadığını ortaya koymuştur. Şans faktörü sayıltısının karşılanıp karşılanmadığı belirsiz kalmıştır. Khi kare istatistikleri yalnız iki parametreli MTK modelinin sınav verisine uyduğunu göstermiştir. Yetenek ve madde parametrelerinin her iki MTK modelinde de değişmezliği yakaladığı bulunmuştur.

MTK ile kestirilen yordama geçerliği klasik toplam puanlarla hesaplanan yordama geçerliğine hem tüm sınav hem de alt test düzeyinde üstün gelmiştir. Alt test düzeyinde okuma alt testinin bölüm İngilizcesi derslerindeki geleceğe ait performansın en iyi belirleyicisi olduğu saptanmıştır.

Anahtar Sözcükler: Madde Tepki Kuramı, Yordama Geçerliği, Yetenek Kestirimi, BÜEYS 2001, BİD 1-2, Madde Bilgi İndeksi, Test Bilgi Eğrisi

ACKNOWLEDGEMENTS

I would like to express my gratitude to my thesis supervisor, Prof Dr. Giray Berberoğlu for his guidance and invaluable suggestions. My gratitude is extended to my colleagues at Başkent University Testing Unit for their help and cooperation. I am also thankful to my colleague Oya Perim Yeğin for her cooperation in all the projects and presentations we have been involved in together.

I would like to thank my mother and father for their ongoing encouragement and support from the beginning of my education life till the completion of this study.

Lastly, I would like to thank my dear wife, Damla Pınar, for her encouragement, help, support, love, and patience for 4 years.

“I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.”

Date: Signature

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURE

CHAPTER I

INTRODUCTION

This study aims at analyzing the predictive power of a fitted IRT model's ability estimations obtained from BUSPE 2001. In this chapter the background of the study, statement of the main and subproblems, the context of the study and the significance of the study are presented respectively.

1.1 Background to the Study

Tests have become an integral part of human life. Especially tests used for educational and occupational purposes influence our lives profoundly. Thus, issues related to their quality have become popular and crucial.

Linn and Gronlund (2000) refer to tests as instruments or systematic procedures for measuring sample behaviors. From another perspective, Croanbach (1990) states that tests are used to assist in decision making like selecting and classifying individuals, evaluating educational or treatment procedures, and accepting or rejecting scientific hypotheses.

Tests may be used for various purposes. According to Aiken (1997), tests may be used to screen applicants for jobs and educational training programs, to classify and place people in educational and employment context, to counsel and guide individuals for educational, vocational and personal counseling purposes, to retain or dismiss, promote and rotate students or employees in educational training programs and in on-the-job situations, to evaluate cognitive, intrapersonal and interpersonal changes due to educational, psychotherapeutic and other behavior intervention programs, and finally to conduct research on changes in behavior overtime and evaluate the effectiveness of new programs or techniques.

When it comes to language testing, tests may be considered in different categories according to their purposes. Alderson, Clapham and Wall (1995) state that language tests tend to fall into the following broad categories: placement, progress, achievement, diagnostic, and proficiency.

Placement tests aim at assessing students' level of language ability so that they can be placed in the appropriate course or class (Alderson, Clapham & Wall, 1995).

Similarly, Linn and Gronlund (2000) point out that placement assessment is concerned with the student's entry performance and focuses on whether students have the knowledge and skills required to start a planned instruction, to what extent students have already mastered the goals of the planned instruction and to what extent the students' interests, work habits and personality characteristics imply that a certain mode of instruction might be better than another.

Progress tests intend to measure the extent to which the students have mastered the material taught in the classroom (Heaton, 1990). They are short testing instruments, which are given regularly to detect how well the students have acquired the language areas or skills, which have just been taught. These tests look back at what students have achieved, therefore they are called progress tests. Encouraging and motivating students is another function of these tests. They also help students to see whether they have achieved short-term objectives or not (Enginarlar, 2002).

Heaton (1990) indicates that achievement tests are based not necessarily on what students have actually learnt or on what has actually been taught, but on what students are presumed to have learnt; in other words achievement tests try to measure the amount of learning in a prescribed domain and should be in line with explicitly stated objectives of a program (Henning, 1987).

Diagnostic tests seek to identify those areas in which a student requires further assistance. These tests can be general and find out whether students have problems with one of the four main language skills; or they can be more specific, seeking to identify problematic areas in a student's grammar (Alderson, Clapham & Wall, 1995).

Enginarlar (2002) defines diagnostic tests as tests that aim to identify the specific strengths and weaknesses in selected areas of language.

Proficiency tests are designed to measure overall control and use of language and communicative skills. These tests are not based on any training program and its syllabus content, they test the readiness of test-takers to perform a task or to follow a course of study in the foreign/second language. Proficiency tests are generally used for specific purposes; for instance, to determine the selection of students that will study in an English-medium school (TOEFL-Test of English as a Foreign Language, the Michigan Test, the British Council Test for studying England) and to judge whether candidates are ready to carry out specific tasks in a work activity when they apply for jobs (Enginarlar, 2002).

Heaton (1990) adds to this by stating that proficiency tests look forward; and that they identify students' language proficiency with reference to particular tasks, which they will be required to perform.

Proficiency tests are designed to test the ability of students with different language training background. Some proficiency tests are intended to show if students have reached a given level of general language ability. Others are aim at showing whether students have sufficient ability to be capable of using a language in some specific area such as medicine, tourism or academic study. Such tests are often referred to as Specific Purposes (SP) tests, the content of which is generally based on a needs assessment of the kinds of language that are needed for a specific purpose (Alderson, Clapham & Wall 1995).

Stating that proficiency tests look forward to test-takers' future performance in a foreign language situation, the issue of the predictive validity of these tests becomes of vital importance.

Hughes (1989) conveys that predictive validity concerns the degree to which a test can predict candidates' future performance. He also emphasizes the necessity of a valid criterion measure against which a proficiency exam is validated. Since other factors than ability in English such as subject knowledge, intelligence, motivation, health, happiness and length of interval could alter a criterion measure's score, a very high validity coefficient is not expected; that is to say, one around 0.4 (only 20 percent agreement) would be acceptable.

The validity coefficient is obtained simply by correlating the raw scores of a proficiency test with the raw scores obtained from the criterion measure. Concerning raw scores of proficiency tests and criterion measures for establishing the predictive validity of the former, reveals the debate of Classical Test Theory (CTT) and Item Response Theory (IRT). What if we used IRT estimations for predictive validity analyses ? Would we be more accurate in assessing test-takers performance by IRT estimations if this proved to be better ? These questions also constitute the core of this study, and answers to them may provide testers with an alternative approach to exploiting test results.

In the related literature IRT and CTT are compared most commonly in favor of IRT. CTT has some shortcomings the most important of which is that test the taker characteristics and test characteristics are inseparable; both can only be interpreted in each other's context. The test-taker characteristic is the ability measured by the test. In CTT this is the raw score of the test, however, this raw score may vary according to the nature of the test. If the test consists of difficult items, the test-taker will have a low ability score; and if the test consists of easy items the same test-taker will have a high ability score. This issue is referred to as test dependency (Hambleton, Swaminathan & Rogers, 1991). The item characteristics in CTT are item difficulty and item discrimination. Item difficulty (p- value) is the proportion of test-takers who respond to an item correctly (Anastasi & Urbina, 1997), and the difficulty of an item may change significantly when high and low ability test-takers take the same test on different occasions. Item discrimination is the degree to which an item discriminates between test-takers with high and low test performance (Anastasi & Urbina, 1997). Similarly, the discrimination index of an item in the same test may vary across different groups of test-takers. For instance, the discrimination index of the same item may be significantly different when it is obtained from a group of homogeneous test-takers and from o group of heterogeneous test-takers on different occasions. These constitute the issue of group dependency. To sum up, when the test-taker context changes the test and item characteristics change and in the same way when the test context changes, the test-taker characteristics change (Hambleton, Swaminathan & Rogers, 1991).

Another shortcoming of CTT is that it is not possible to compare test-takers scores across different forms of a test or even different sections of a test without applying sophisticated equation procedures (Hambleton & Oakland, 1995).

The assumption of equal errors of measurement for all examinees is the third shortcoming of CTT, because the error of measurement on a test may be different when it is given to different test-takers. This occurs especially when a difficult test is administered to high and low ability test-takers (Hambleton, Swaminathan & Rogers, 1991).

Finally, CTT does not provide the tester with any information about how a testee might perform on a given item. This might especially be crucial when a test constructor wants to design tests with particular characteristics for certain populations of testees (Hambleton, Swaminathan & Rogers, 1991).

Contrary to the shortcomings of CTT, Hambleton, Swaminathan and Rogers (1991) point out the desired features of an alternative test theory, the IRT, which includes:

- Item characteristics that are not group dependent
- Test characteristics that are not group dependent

which is:

- A model that is expressed at the item level rather than at the test level
- A model that does not require strictly parallel tests for assessing reliability
- A model that provides a measure of precision for each ability score

Therefore, the predictive validity coefficient calculated for a test by using the ability estimates of a fitted IRT model could be more precise and not group dependent. However, an ability estimates for both the criterion variables and the predictor variables may not be available, then the predictive validity coefficient may be group dependent to some extent. Yet it can be expected that the IRT estimated predictive validity of a measure may be superior to its classical predictive validity due to the above mentioned advantages of IRT over CTT.

1.2 Statement of the Purposes

The main purpose of this study is to analyze the predictive validity of the BUSPE 2001 by using the estimations of either the one or two parameter IRT model.

This study follows two stages. In the first stage, the fit of the one and two parameter IRT models to BUSPE 2001 is investigated, which is a prerequisite analysis for the IRT estimated predictive validity analyses. This fit analysis requires a number of assumptions to be checked and invariance analyses to be conducted, because to obtain good IRT estimates the test data needs to meet certain conditions. For these prerequisite analyses the following questions are investigated in the data set.

1.1 Does the BUSPE 2001 meet the assumptions of IRT ?

1.1.2 Is the BUSPE 2001 a Non Speeded test ?

1.1.3 Are the items of BUSPE 2001 locally independent ?

1.1.4 Does the BUSPE 2001 meet the assumption of equal item discrimination indices of the one parameter model ?

1.1.5 Does the BUSPE 2001 meet the minimal guessing assumption of one and two parameter IRT model ?

1.1.1 Does the BUSPE 2001 measure a unidimensional trait ?

1.2 How well do the simulated test results of the one and two parameter model predict actual test results ?

1.2.1 How well does the observed distribution of the BUSPE 2001's scores fit the theoretical distribution of the one and two parameter IRT model ?

1.2.2 How strong is the relationship between ability estimates of the one and two parameter IRT models, and the actual total scores of the BUSPE 2001 ?

1.3 Are the item parameters estimates of the one and two parameter IRT models invariant across different samples of examinees ?

1.3.1 Are the item parameters estimated by the one and two parameter IRT models invariant across high and low scoring groups ?

1.4 Are the ability parameter estimates of the one and two parameter IRT models invariant across different item sets of the BUSPE 2001 ?

1.4.1 Are the ability parameters estimated by the one and two parameter IRT models invariant across sets of easy and difficult items of the BUSPE 2001 ?

1.4.2 Are the ability parameters estimated by the one and two parameter IRT models invariant across sets of odd and even items of the BUSPE 2001 ?

In the second stage, the predictive validity of BUSPE 2001 is examined in the light of the analyses conducted in the first stage. In fact the first stage serves as a preliminary analysis for the second stage which is of primary concern to this study.

1.3 Statement of the Main and Subproblems

The following research questions state the problems that make up the core of this study:

1.1 How well does the BUSPE 2001 predict on departmental English courses (DEC) at Başkent University ?

1.1.1 How well do the total scores of the BUSPE 2001 predict on first and second semester DEC grades in the freshmen year (DEC 1-2) ?

1.1.2 How well do ability estimates of the fitted IRT model predict on DEC 1 and DEC 2 ?

1.1.3 How well do the ability estimates obtained by using the fifty items providing highest information in the fitted IRT model, predict on DEC 1 and DEC 2 ?

1.1.4 How well do the total scores of the fifty items providing highest information in the fitted IRT model predict on DEC 1 and DEC 2 ?

1.1.5 How well do the total scores of the thirty items providing highest information in the fitted IRT model predict on DEC 1 and DEC 2 ?

1.1.6 How does reducing the number of items to fifty and thirty affect content validity, construct validity and reliability ?

1.1.7 How well do the total scores of the grammar section predict on DEC 1 DEC 2 ?

1.1.8 How well do the ability estimates obtained by using the grammar section in the fitted IRT model, predict on DEC 1 and DEC 2 ?

1.1.9 How well do the total scores of the reading section predict on DEC 1 and DEC 2 ?

1.1.10 How well do the ability estimates obtained by using the reading section in the fitted IRT model, predict on DEC 1 and DEC 2 ?

1.1.11 How well do the total scores of the vocabulary section predict on DEC 1 and DEC 2 ?

1.1.12 How well do the ability estimates obtained by using the vocabulary section in the fitted IRT model, predict on DEC 1 and DEC 2 ?

1.4 Context of the Study

Başkent University (BU) is a Turkish medium school, however, each student registering is required to be at a certain proficiency level of English to start as freshman. This proficiency level is first diagnosed by a placement exam and then tested by a proficiency exam. Firstly, students take the placement exam, those who pass can take the proficiency exam, the others start the preparatory school of BU at C-level. Students passing the proficiency exam start as freshman. The students who fail to pass the proficiency exam start the preparatory school at B-level. For both exams the passing score is 60. During the preparatory year students are tested on eight achievement tests. They have to reach an overall of 60 to be able to take the proficiency exam. Those students who get a minimum score of 60 pass to freshmen. The others can register to the BU summer school program. All students who attend the summer school are given an extra chance of taking a proficiency exam. If they get a score of 60 and above, they pass to freshmen. The others have a final chance and take another proficiency exam with the students who did not attend the summer school course and with those have just enrolled the university and those who failed in the preparatory school due to unattendance. This final proficiency exam is administered in September.

At Başkent University proficiency exams are administered to check whether students have adequate capacity to attend and succeed in departmental English courses.

1.5 Significance of the Study

The significance of this study can be explained from two points of views. This is because it is not only a predictive validation study, but also a study analyzing its fit to the one and two parameter IRT models.

Considering both views, the following reasons make this study important for language testing practices both at BU and other institutions.

1. Adaptive tests require items that provide precise information about test-takers' ability and that are at the difficulty level appropriate to the test-taker's ability determined by preceding items. Thus, this study will contribute to adaptive testing practices, as IRT models provide the necessary information for adaptive tests in the most reliable way.

2. In constructing item banks, items are selected according to their difficulty, and to how well they discriminate among different test-takers with different abilities. IRT models provide invariant difficulty and discrimination indices, and item characteristic curves (ICCs), which provide information about the ability level the items are best geared to and the effectiveness of the items in general. Hence, this study will provide precious information for item banking efforts, especially at BU.

3. This study will help testers to evaluate how different items of a language proficiency test function, especially by providing ICCs for each item, and therefore will provide clues for reshaping such tests.

4. This study will also compare CTT and IRT indices in terms of preciseness and usefulness for predictive validation. Thus, it might suggest an alternative approach that is more powerful for predictive validation studies. Consequently, this might contribute to utilization of test scores, that is to say, the results of this study might reveal that IRT ability estimates of a test might be more reliable and valid than raw scores for any decisions considering the test results.

5. Paper and pencil language proficiency tests usually contain separate sections that intend to test grammar, reading, vocabulary and writing. They may have different weightings in different tests. This study will either justify the present weightings of the BUSPE 2001 or provide evidence for different weightings, and in this sense tips for constructing language proficiency tests will be given.

6. The results of the predictive validation of this study will also provide useful ideas for other test validation studies like convergent and divergent validation.

7. Taking language testing into consideration, this study will contribute to validating and constructing not only proficiency exams, but also achievement, progress and diagnostic tests.

CHAPTER II

REVIEW OF THE LITERATURE

In this chapter, the related literature survey about IRT and predictive validity of tests, mainly in the context of language testing are presented.

2.1 Item Response Theory

When tests are analyzed it is impossible to separate the test-takers' characteristics from the tests' own characteristics, this means that the results of analyses are only valid for the samples they were carried out on. Concerning language proficiency exams, the results of analyses will not be valid for students at different proficiency levels. Thus, a proficiency test's difficulty may not be fixed. The test may be difficult for one group of test-takers who are not very proficient, while it may be easy at the same time for another group who are highly proficient. Therefore, it may be difficult to compare students who have taken different tests, or to compare items which have been tried out on different groups of test-takers. In order to cope with this problem, measurement using Item Response Theory was developed. Thanks to IRT, performance of test-takers who have taken different tests can be compared or the same item analysis can be applied to groups of test-takers with different levels of language proficiency by developing an item difficulty scale independent of the sample on which the items were tested. Hence, it is theoretically not necessary to conduct both tests for the same group of test-takers. Provided that a few identical 'anchor' items are included in the two versions of the test, each version can be trailed on a different group, and the two versions can be equated by these anchor items (Alderson, Clapham & Wall, 1995).

Hambleton, Swaminathan and Rogers (1991) state that IRT has two claims the first is that performance of an examinee on a test item is predictable by a set factors called traits, latent traits or abilities, these latent traits can not be directly measured. A latent trait is generally called the ability measured by the tests, it is represented by the symbol 'θ' and the total score is the initial estimate of the ability (Anastasi & Urbina, 1997). The second is that the relationship between test-takers' item performance can be described by a monotonistically increasing function called an item characteristic function or item characteristic curve (ICC). According to this function as the level of the trait or ability increases, the probability of a correct response to an item increases. (Hambleton, Swaminathan & Rogers, 1991).

There are many possible item response models, the difference of which is the mathematical form of the item characteristic function and/or the number of parameters specified in the model. There is at least one parameter describing the item and at least one parameter describing the test-taker in every IRT model. The initial step in any IRT practice is to estimate these parameters (Hambleton, Swaminathan & Rogers, 1991).

These parameters, which usually describe the ICC, are the “b”, “a” and the “c” parameters where “b” refers to the difficulty, “a” to the discrimination and “c” to the lower asymptote parameter (Weiss & Yoes, 1991). Information is higher when the b value is close to “θ” than when the “b” value is quite different from “θ”, information is generally higher when the “a” parameter is high, and information increases as the “c” parameter approaches zero (Hambleton, Swaminathan & Rogers, 1991).

Item information functions may contribute much to test development and item evaluation, because they show how the items contribute to ability estimation. Item information functions’ utility is bound to the fit of ICCs. If the fit of the ICCs to the data is poor, the corresponding item statistics and item information functions may not be accurate (Hambleton, Swaminathan & Rogers, 1991). The ICCs of the most and least informative items (Item 12 & 94) of BUSPE 2001 are displayed as examples in Appendix 1 to domenstrate how ICCs of good and poor items look like.

The sum of the item information functions at “θ” makes up the information provided by a test at “θ”. Thus, how much individual test items contribute to test information functions can be determined without knowledge of the other items in the test (Hambleton, Swaminathan & Rogers, 1991).

The most common IRT models are the one, two and three parameter logistic models:

The one parameter (Rasch) Model is the simplest of three models. It has fewer requirements than the other two. A minimum of 100 test-takers is usually considered adequate. A smaller sample results in a higher measurement error. However, this model is limited in scope, because it takes only ability and item difficulty into consideration (Alderson, Clapham & Wall, 1995). It assumes that there is no guessing and the "a" (discrimination) parameters are homogeneous.

The two parameter model operates one step beyond the one parameter model. It also takes into account the item discrimination index (the a parameter). It assumes that the "c" (lower asymptote) parameter is zero. It requires a sample of at least 200 students (Alderson, Clapham & Wall, 1995).

The three parameter model is the most sophisticated model. In addition what the two parameter model does, it also takes guessing (the "c" parameter) into account. It requires a data set of 1000 test-takers (Alderson, Clapham & Wall, 1995).

IRT is based on strong assumptions about the nature of the test data. One assumption is unidimensionality which is defined in terms of the statistical dependence among items. If the statistical dependence of the items in a test can be explained by a single latent trait, it can be considered as unidimensional (Crocker & Algina, 1986). For Anastasi and Urbina (1997), unidimensionality assumption is sufficiently met when test performance depends on a single trait, even though other traits slightly affect performance. Unidimensionality can be checked by a factor analysis and is required to be met by all three IRT models (Hambleton, Swaminathan & Rogers, 1991).

Another assumption is local independence. It is enhanced when the probability of a correct response of a test-taker to an item is not affected by responses to other items in the test. In other words, local independence means that test-taker responses to any pairs of items are uncorrelated when the ability influencing test performance is kept constant. Thus the inter-item correlations' means obtained from different ability groups should be close to zero (Hambleton, Swaminathan & Rogers, 1991).

The one parameter model requires the assumption of equal discrimination indices to be met. According to this assumption the item discrimination indices of a test obtained from a standard item analysis must be reasonably homogeneous, so that the one parameter model may be viable (Hambleton, Swaminathan & Rogers, 1991).

Nonspeeded test administration is another assumption of IRT. To meet this assumption all test-takers must attempt to answer all the items in the test. Therefore, if a test-taker knows the correct answer he/she will answer it correctly, while he/she will probably answer the item incorrectly if he/she does not know the correct answer. To check nonspeededness, the variance of number of omitted items could be compared to the variance number items answered incorrectly.

Minimal guessing is the assumption of the one and two parameter models. In order to check this assumption the performance of low ability students on the most difficult items can be checked and if performance levels are close to zero, the assumption is said to be met (Hambleton, Swaminathan & Rogers, 1991).

The main claim of IRT models is that test characteristics and test-taker characteristics are invariant of each other. In other words, sample free item parameter estimates and item free ability estimates are obtained by IRT models (Hambleton & Oakland, 1995). This assumption needs to be checked as well. For that, item parameter estimates can be obtained from different samples and than compared, and ability parameters can be obtained with different sets of items and compared afterwards. Chi square statistics and checking model predictions of actual and simulated test results, which is done by correlating ability estimates of the IRT models with raw scores, also assists in deciding on the model data fit (Hambleton, Swaminathan & Rogers, 1991)

To date numerous test analysis studies involving IRT applications have been conducted in Turkey and abroad. For instance, Kılıç (1999) analyzed the fit of the one, two and three parameter IRT models to the 1993 Student

Selection Test's Mathematics, Natural Sciences, Turkish and Social Sciences subtests. The data of 2121 test-takers were analyzed. It was found out that none of the subtests met the assumption of equal discrimination indices. All subtests except the Turkish subtest was found to be speeded to some extent. Invariance of the ability estimates was found to be most invariant across different sets of items compared to the other subtests. The three parameter model item discrimination parameter estimates of the Mathematics and Natural Sciences subtests were less invariant across different samples of examinees than the two parameter model. Nevertheless, in the Turkish and Social Science subtests it was observed that the two and three parameter models' items discrimination indices were highly invariant across different samples of test-takers. In addition Chi square statistics results indicated that the fit of the three parameter model to the SST was better.

Özkurt (2002) also conducted a fit analysis. She examined the fit of the one, two and the three parameter IRT models to an English Proficiency Test administered at a state university in 2000. The test data of 361 students was used for the study. The findings indicated that the assumptions of unidimensionality, nonspeededness and local independence were met. However, the assumption of invariance of item and ability parameters was not met. The number of misfit items was found to be lowest in the two parameter IRT model. The two parameter IRT model was found to the best model to fit the English Proficiency Test data. The test data of 468 students was subject to this study.

Similarly, Karataş (2001) carried out a study to investigate the fit of the one-, two- and three-parameter models of IRT to the English Proficiency Test administered at Erciyes University in 1999. The findings of this study revealed that the test was unidimensional, the items were locally independent and that there was minimal guessing. The test was found to be nonspeeded. The discrimination indices of the test's items were not homogeneous. All three IRT models' item parameters were invariant across different groups of examinees. However, the two-parameter IRT model seemed to provide more invariant item difficulty parameters. All three IRT models had invariant ability estimates, in fact the one and two parameter IRT models' were found out to be yielding more invariant ability estimates than the three parameter IRT model. The chi-square statistics demonstrated that the fit of the one parameter model was poorer compared to the two and three parameter IRT models.

Stage (1998) compared item statistics of the 1997 SweSAT word subtest from the CTT framework and those from the IRT framework and examined the stability from pretest to regular test of the two sets of items statistics for two groups, males and females. The study revealed that the three methods which were using CTT, Mantel Haenszel and IRT appeared to give similar results, and that there was higher agreement between MH and CTT regarding differential item functioning. The comparison with IRT was made solely with the estimated b parameter of each item and Stage (1998) reports that the advantage of IRT was that information was given along the whole ability continuum and not for a single point as with CTT.

Fan (1998) conducted a study to examine the behavior of item and person statistics obtained from the CTT and IRT measurement frameworks. The study's focus was on two issues: a) How comparable are the item and person statistics from CTT from those of IRT ? b) How invariant are the CTT item statistics, respectively ?. Random samples of 1000 examinees were chosen from a statewide assessment program. The test item pool consisted of a math tests with 60 and reading test with 48 dichotomous items. The findings of the study were as follows.

1. The ability estimates obtained from CTT were highly comparable with those of the three IRT models.
2. Item difficulty indices of CTT were very comparable with all IRT models especially the one parameter model.
3. Item discrimination indices of CTT compared to item difficulty indices were less comparable with those from IRT.
4. Both CTT and IRT item difficulty indices were highly invariant across different samples, and there was no great difference between the two.
5. Both the CTT and IRT item discrimination indices were not as invariant as the difficult indices. The degree of invariance of CTT

item discrimination indices was highly comparable with that of IRT item discrimination estimates.

2.2 Predictive Validity of Language Tests

In many cases the test user wants to draw inferences from test scores in order to analyze behavior on some performance criterion which cannot be directly measured by a test. In such cases, decision makers must have evidence that there is a relationship between test score and criterion performance before using the test scores to make decisions like admission or hiring. Such kind of evidence is obtained from a criterion-related validation study (Crocker & Algina, 1986).

In a more detailed way, Crocker and Algina (1986) suggest that the general design of a criterion-related validation study follows these steps: First, determining a suitable criterion behavior and a method for measuring it, then determining an appropriate sample of examinees representative of those for whom the test will finally be used, after that administering the test and recording each examinee's score, next obtaining a measure of performance on the criterion for each examinee when the criterion data are available and finally determining how strong the relationship between test scores and criterion performance is.

According to Crocker and Algina (1986), there are types of criterion-related validation: concurrent and predictive: Where concurrent validity refers to the relationship between test scores and criterion performance measured at the time the test was given and predictive validity refers to the degree the test scores predict criterion performance that will be measured in the future. Enginarlar (2002), states that one should correlate the scores of the test that will be validated with a future criterion performance in order to establish predictive validity.

To exemplify, the Scholastic Aptitude Test (SAT) may have a degree of predictive validity with respect to college grade point average (CGPA), because SAT scores correlate about .40 with CGPA. This justifies the use of SAT scores in making admission decisions. College admission directors would prefer admitting those students who will be academically successful in college and CGPA can prove this academic success. Thanks to the demonstrated relationship between SAT scores and CGPA, the use of SAT scores in admission decisions is to some extent justified for drawing inferences about examinees' future performance (Crocker and Algina, 1986).

Anastasi and Urbina (1997) claim that the term "prediction" can be used both in the broader sense and in the more limited sense. That is to say; both to refer to prediction from the test to any criterion situation and to prediction over a time interval. The term "predictive validity" is used for the more limited sense of prediction. The information provided by predictive validation could be relevant to tests used in the selection and classification of personnel. Hiring job applicants, selecting students for admission to college or professional schools, and assigning military personnel to occupational training programs are decisions taken through tests that require information about their predictive validity (Anastasi & Urbina, 1997). Heaton (1990), exemplifies predictive validity with the correlation between an English test that will be administered to engineers who will take civil engineering courses and their measured performances on these courses.

In the context of language testing Alderson, Clapham and Wall (1995) express that predictive validation is most common with proficiency tests: tests which are designed to predict how well somebody will perform in the future. The simplest form of predictive validation is to give students a test, and then at some time in the future give another test of the ability that the first test claimed to predict. A common use for a proficiency test like IELTS or the TOEFL is to identify students who might have problems when studying in an English medium setting because of weaknesses in their English. Predictive validation involves giving students the IELTS test before they leave their home country for overseas study, and when they are in the host study setting and have settled down, giving them a test of their use of English in that study setting. A high correlation between the two scores indicates a high degree of predictive validity for the IELTS test (Alderson, Clapham & Wall, 1995).

However, Alderson, Clapham and Wall (1995) convey the problem of "truncated samples". Not all students taking the IELTS test are able to travel to the overseas, some are excluded because of poor test performance. This is known as the “truncated sample problem”. One can only use a part of the original test population in the validation – in this case, those who can be used will be the better students. The effect of using truncated samples is not well known for such tests, yet the spread of students’ scores tend to decrease, and to depress the predictive validity coefficient. If all the students were allowed to enter English-medium education instead of only the best students, the correlation between the two tests would be higher. Secondly, it is likely that in this example, the language proficiency of the students might improve between the first and second occasions, especially once they have arrived in the host country. This has also the effect of depressing the predictive validity coefficient. Thirdly, as with the concurrent validity, it is unlikely that a suitable external measure of the students’ ability to use English in the study setting will be available, unless it is another version of the original test. The latter problem is a threat to many predictive validity studies, because one needs a good measure of the skill he/she is trying to predict. Some validation studies of proficiency tests exploit the class of degree of Grade Point Average (GPA) that the students get at the end of their studies. However, in such studies the use of truncated samples is not the only threat, the results of any correlations are also obscured by the fact that the class of degree/GPA reflects not only language ability, it also reflects academic ability, subject knowledge, perseverance, study skills, adaptability to the host culture and context, and many other variables (Alderson, Clapham & Wall, 1995).

Alderson, Clapham and Wall (1995) put forward that besides GPA other measures can also be used in predictive validation. According to them, one might attempt, for example, to gather the opinions of those who come into regular contact with the students. The test validator might ask instructors to rate the students who have taken the test on their language abilities: their writing ability, their oral communicative abilities, etc. Here again, there may be only a truncated sample available. There will also be the problem that many instructors are not able to give a useful opinion about their students' language abilities until probably the end of the first terms. Then the students might already have had the opportunity to improve their language. The resulting correlations are very difficult to interpret.

Alderson, Clapham and Wall (1995) also state that in any predictive validity study a high correlation should not be expected. They mention that a correlation coefficient of +.30 would be satisfactory for many researchers in validating a test.

When working on the predictive validity of a test it is crucial to know what purpose the test will be used for. It might be of no sense to bear predictive validity consideration as it is in the case of the TOEFL when it is being employed for decision making purposes only to ensure that students who are proficient in English are allowed to enroll in English speaking universities, but instead is not being used as a standard for admission with the aim of predicting future academic achievement (Simner, 1999).

Pack (1968, as cited in Hale, Stansfield & Duran, 1984) found that TOEFL scores were significantly related to the grade obtained in the first English course taken, and that they were neither related to grades obtained in subsequent English courses and nor related to the probability that an examinee would graduate.

Schrader and Pitcher (1970, cited in as cited in Hale, Stansfield & Duran, 1984) reported that after an eight-week summer university orientation program given in English, students' scores on the TOEFL itself increased. If the cutoff of 600 is strictly adhered to, applicants who are rejected due to scores of 580 to 590 might well increase their scores to 600 or more if given the opportunity to improve their English after arriving on campus and before starting classes. Therefore, it would seem that any attempt to justify the use of the TOEFL as a means of predicting a university applicant's future command of English based solely on evidence from concurrent validity studies with the TOEFL must be viewed with some suspicion.

Dooey's research (1999) studied the predictive validity of the IELTS Test. She tested whether the IELTS test is a predictor of academic success or not. Her findings did not provide conclusive evidence about the validity of IELTS as a predictor of academic success. However, the reading section had the highest correlation with academic success ratings.

De Noble, Jung and Ehrlich, (1999) conducted a study in which they tried to develop a measure of entrepreneurial self-efficacy that would have high validity in predicting entrepreneurial action. For this, they conducted a factor analysis for their measure and detected the distinct factors. After that, they formed sub tests for those factors and correlated their scores with their actual entrepreneurial action to find out which factor's sub test would have higher predictive validity. This study did not deal on the item, but subtest level in determining a measure's predictive validity.

Prapphal's study (1990) examined the predictive validity of different types of language tests on academic achievement in General English and EAP courses. For the study 264 science students who had taken the national English entrance examination in Thailand in 1982 were selected. A hundred and thirty-nine of these subjects took freshman general English courses and 125 subjects took English for Academic Purposes courses. Multiple choice cloze and matching cloze tests, representing global knowledge of English and cognitive processing abilities, were compared with a traditional reading comprehension test which represents a less synthesized knowledge of linguistic elements. It was found that the test format, in this study the cloze test, may be significant in predicting future academic achievement, and the content of language tests may play a role in academic achievement for each type of language program.

Another study of Prapphal (1990) tried to show the direct and indirect relationships between subskills of General English and EAP tests. All language subskills, regardless of content, were found to be significantly related. Therefore, it was concluded that there might be a transfer of subskills from one content to another. It was also found that the nature of the transfer of language subskills might be direct or indirect. However, a hierarchical relationship from General English to English for Academic Purposes was not confirmed by this study. The transfer of language subskills across content was reported to have probably occurred within each subskill. Students who had mastered subskills in vocabulary, structure, and reading in General English might have transferred these subskills to English for Academic Purposes. However, this observation did not apply to the writing subskill in both tests.

Prapphal (1990) also conducted a study with 100 first year students. He examined the underlying relationships between General English and EAP tests. It was found that EAP tests may predict achievement in EAP programs better than General English tests; the formats associated with each discipline tended to predict academic success in science better than those that were not related to a specific discipline; and there was a common factor shared by the EAP tests, General English tests, and knowledge of the subject matter represented by student grade point average.

Patitas (1989) analyzed the predictive validity of the Khon Kaen University Entrance Examination Tests for non-Science program which consisted of 10 tests: General Subject II, Mathematics I, English I - II - III, Thai Language I- II, Social Studies I - II, French, Fundamental Mathematics, Fundamental Thai Language, Fundamental Science, and Fundamental Social Studies. 8004 of the 19,446 total test-takers were selected for the study. The results of the entrance exam and the first semester grade point average (GPA) of 123 first year students from 2 faculties who passed the entrance exam were used for analyzing the predictive validity of the tests. The predictive validity coefficients of French and Fundamental Mathematics were found to be significant.

To sum up, it is observed that predictive validity of measures are investigated by simply correlating the raw scores of criterion measures with the raw scores of predictor measures and validity coefficients between .3 and .4 are usually considered as acceptable. Language proficiency tests tend to provide a good fit to the two parameter IRT model. Consequently, it seems wise to analyze the predictive validity of BUSPE 2001 through the ability estimates of the two parameter IRT model and an improvement in the predictive validity coefficient could be expected.

CHAPTER III

METHOD OF THE STUDY

In this chapter the methodological procedures are presented. The overall research design, the subjects, the instruments and the data analysis techniques constitute the topics. In the first section the overall reseach design is summarized. In the subject section, some descriptive data of the students who are subject to this study are explained. In the instrument section, the properties of BUSPE 2001 including its subtests and item types, and the DEC grades that are used in the predictive validity study are expounded. In the data analysis section, the procedures undergone for checking model assumptions, chi square statistics and assessing invariance, which serve as preliminary analyses for analyzing the predictive validity of BUSPE 2001 are described.

3.1 Overall Research Design

The main purpose of this study is to analyze the predictive validity of the BUSPE 2001 by using the estimations of a fitted IRT model. First the fit of one and two parameter IRT models to BUSPE 2001 was investigated. By using the test data of all students (727) who took BUSPE 2001. After the fit analyses the predictive power of the fitted IRT model's estimates were analyzed in comparison to classical raw scores. All DEC 1-2 scores were used as criteria for the predictive validity analyses. Three hundred and twenty-five students who passed BUSPE 2001 had a DEC 1 score and 184 had a DEC 2 score.

3.2 Population and the Sample

The subjects of this study were all students who took BUSPE 2001. That is 727 students. 543 of them were new registered students and 184 of them were students who had failed in the preparatory school in the 2000-2001 academic year. Among the students who passed BUSPE 2001 not everybody had a DEC 1 or DEC 2 score, because some departments did not have English courses in their curricula in these semesters or some of the students chose either not to continue or not to take the course. Three hundred and twenty-five of the students who passed BUSPE 2001 had a DEC 1 grade and 184 had a DEC 2 grade.

3.3 Data Collection Instrument

In this study, the data used were obtained from the test data of BUSPE 2001 and DEC 1-2 grades.

BUSPE 2001 like all other proficiency exams of Başkent University consisted of 100 items all of which were multiple choice with four alternatives. These exams had 3 sections: Grammar, reading, vocabulary respectively. The grammar section consisted of the first 40 items: Items 1 to 15 made up the modified cloze test, items 16 to 35 were discrete point items, and items 36 to 40 were spot the mistake type. Items 41 to 81 constitute the reading section: Items 41 to 45 are sentence completion items, items 46 to 50 were paragraph completion items, items 51 to 60 were sentence completion and reference type items of the first reading text, items 61 to 70 were sentence completion and reference type items of the second reading text, and items 71 to 80 were sentence completion and reference type items of the third reading text. Items 80 to 100 formed the vocabulary section all of which were sentential level fill in the blanks type multiple choice items.

The criteria used to establish the predictive validity of the BUSPE 2001 were the DEC 1-2 grades. The DEC 1-2 grades were not obtained from single measures, but from several sources. These grades were obtained by adding the following weightings of four grades: 30 % of a midterm exam

(Achievement exam testing grammar, reading comprehension, vocabulary and writing.) grade, 10 % of a project exam (Alternative assessment testing reading comprehension or speaking skills depending on the department) grade, 10% of a teacher evaluation (Evaluation of the class teacher according to 4 criteria: Participation, attendance, homework, preparation) grade, 50% of final exam (Achievement exam testing grammar, reading comprehension, vocabulary and writing) grade.

3.4 Data Analysis Procedures

For the SPSS part of the study, the data obtained from the optic forms of the BUSPE 2001 were coded dichotomously on the SPSS processor as 0 for incorrect and 1 for correct responses. Then the data was converted into an appropriate format to be run with BILOG (Mislevy & Bock, 1986).

3.4.1 Preliminary Analysis

Descriptive statistics including measures of central tendency (mean, mode, median) and measures of variation (standard deviation, variance, skewness, kurtosis, range), minimum-maximum score, percentiles and frequency distribution with a normal curve were obtained to demonstrate an overall picture of the proficiency exam results.

Classical item analysis was conducted and item difficulty (Item means) and item discrimination (Corrected-item total correlation) indices for each item were obtained to assess how each item of BUSPE 2001 functioned.

The alpha coefficient, which Green, Salkino and Akey (1997) regard as the most appropriate index for reliability of dichotomously scored items was computed for the BUSPE 2001. Each of the 100 items was correlated with the total proficiency score by deleting the item (alpha if item deleted) to see how each item contributes to the internal consistency of BUSPE 2001.

3.4.2 Goodness of Fit Analyses

3.4.2.1 Checking Model Assumptions

The dimensionality of BUSPE 2001 was assessed by a Principal Component Analysis. The eigenvalues obtained were displayed in a scree plot to make a final decision about the number of constructs measured by BUSPE 2001.

To check whether the items of BUSPE 2001 were locally independent the mean score of inter-item correlations were obtained by using the whole group of students and restricted range ability groups. The restricted range ability groups were students in the first quartile with total scores of 40 and below, and student in the fourth quartile with total scores of 68 and above. If these mean scores were close to zero local independence would be enhanced.

The item discrimination indices obtained by classical item analysis were reviewed and plotted in a histogram to observe whether the item discrimination indices were equal. Equal item discrimination indices would make the use of the one parameter viable.

In order to examine whether guessing was a factor affecting the students total test scores the most difficult items were selected and the performance of the low ability students on these difficult items was observed. The low ability students were those in the first quartile. Mean scores close to zero would imply that guessing is not an influential factor considering total test scores, otherwise it would be wise to consider the three parameter model.

Speededness of BUSPE 2001 was analyzed by calculating the ratio of the variance of omitted items to the variance of items answered incorrectly. A ratio close to zero would imply that the test is speeded and none of the three IRT models were viable.

After checking the model assumptions, the ability and item parameter estimates, and chi square goodness of statistics for the one and two parameter models were obtained by using the BILOG program.

3.4.2.2 Checking Model Predictions of Actual and Simulated Test Results

To check how well the observed distribution of BUSPE 2001's scores fit the theoretical distribution of the one and two parameter IRT models, chi square goodness of fit statistics of BUSPE 2001 generated by the BILOG program were used to determine whether the one or two parameter IRT model fit the data better. An insignficant result at alpha level 0.05 would indicate that the model fits the data. Moreover, the test information curves obtained for both the one and two parameter IRT models were analyzed to find out how much and how accurate information each IRT model provided.

So as to examine model predictions of actual and simulated test results, the students' ability estimates calculated for the one and two parameter IRT models were correlated with total test scores. A strong relationship with scatter around the test characteristic curve is expected. The strongest relationship is expected with the IRT model that makes the most accurate predictions.

3.4.2.3 Checking Expected Model Features

The item and ability estimates of IRT modals are expected to be invariant.

The invariance of the ability parameters of the one and two parameter IRT models was checked across different samples of students. Firstly, ability estimates were obtained by using the hardest and easiest fifty items of the test, then they were correlated and the correlation was displayed in a scatter diagram. The same procedure was also applied for odd and even items. The higher the correlation, the more invariant would be the ability estimates.

The invariance of "a" parameters of the one parameter IRT model, and of the "a" and "b" parameters of the two parameter IRT model were checked across different groups of students. Firstly, the students were divided into two groups as high and low performers according to their total test scores. Then the item parameters of the one and two parameter IRT models obtained by using these groups were correlated and displayed in a scatter diagram. High correlations would prove invariance of the item parameter estimates.

3.4.3 Predictive Validity Analyses

The predictive validity of the BUSPE 2001 was analyzed from different perspectives; by using both raw scores and IRT estimations for both the total test and its subtests.

First the total scores of the BUSPE 2001 were correlated with the DEC 1-2 grades to analyze its common predictive validity. After that, the ability estimates obtained from the fitted IRT model were correlated with the DEC 1-2 grades and consequently, the results were compared.

To analyze how the items of BUSPE 2001 that provide high information in the fitted IRT model function in predicting DEC 1-2 grades, fifty items providing highest information were determined by checking ICCs and item information indices. Secondly, the total scores of these fifty items were correlated with DEC 1-2 grades, then by using these fifty items new ability estimates were obtained in the fitted IRT model and correlated with DEC 1-2 grades as well. Lastly, both correlations were compared.

The above mentioned procedure was also applied with thirty items providing highest information in the fitted IRT model.

The effect of choosing fifty items providing high information on the content and construct validity of BUSPE 2001 was also analyzed by checking the type of items that were excluded. In addition, the reliability of these sets of items was investigated as well. Next, the test information curves of these sets of items obtained in the two parameter IRT model were analyzed.

The total scores of the grammar, reading and vocabulary subtests were correlated respectively with DEC 1-2 grades to find out the predictive power of each subtest. After that, by using each subtest separately ability estimates were obtained in the fitted IRT model and respectively correlated with DEC 1-2 grades. Consequently, correlation coefficients were compared.

CHAPTER IV

RESULTS OF THE STUDY

In this chapter the results of the study are presented under three headings. Firstly preliminary analyses, secondly the results of the goodness of fit analyses consisting of three subheadings which are checking model assumptions, checking model predictions of actual and simulated test results, and checking expected model features, thirdly predictive validity analyses.

4.1 Preliminary Analyses

The descriptive statistics found are presented in Table 4.1.1. The mean of BUSPE 2001's total scores was found to be 53.96, which is not very high considering the cut score of 60, the median was 54, the mode was 65 representing the most frequently observed score. The standard deviation was 17.79 and the variance was 316.57 which indicate a large and desirable distribution. The skewness value was .015 showing that the test was not skewed, kurtosis was -.819, moreover, the median and mean were close and the score 45 looked like a second mode, therefore the distribution resembled a bimodal. The range was 89, which means that the difference between the highest (96) and lowest (7) score was quite high. The score on the 25th, 50th, 75th percentiles were 40, 54, 68 respectively.

Table 4.1.1 Descriptive Statistics of BUSPE 2001

NUMBER OF ITEMS		100
NUMBER OF STUDENTS		727
MEAN		53.96
MODE		65
MEDIAN		54
STANDARD DEVIATION		17.79
VARIANCE		316.57
SKEWNESS		.015
KURTOSIS		-.819
RANGE		89
MINIMUM SCORE		7
MAXIMUM SCORE		96
PERCENTILES	25	40
	50	54
	75	68
ALPHA		.94
MEAN DIFFICULTY (p)		.5396
MEAN DISCRIMINATION (r)		.3555

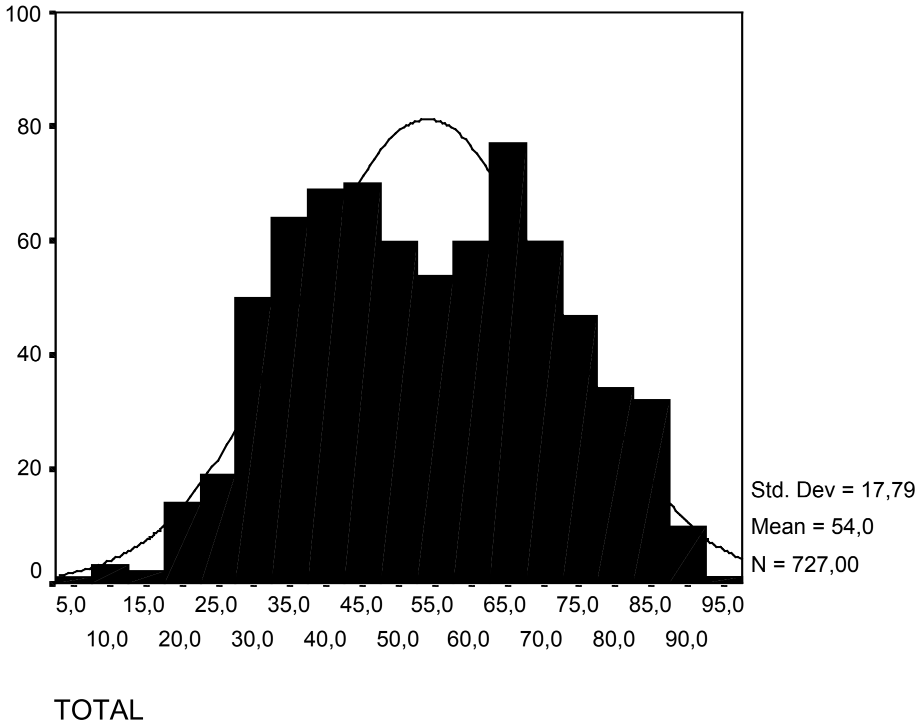

Figure 4.1.1 Frequency Distribution of Total Scores of BUSPE 2001

The item difficulty and "p" and item discrimination indices "r" were obtained through classical item analysis (Appendix 2). The item difficulty indices ranged from .1348 to .9574. The mean of item difficulties was .5396 for the whole test. The means of item difficulties were .5068 for grammar subtest, .608 for the reading subtest and .4726 for the vocabulary subtest. The most difficult subtest seemed to be the vocabulary subtest followed by the grammar and reading subtests respectively. The most difficult items had items difficulties of ,1348 (item 86), .1967 (item 33), .2077 (item 87), .2160 (item 21), .2682 (item 5), .2889 (item 19), .2930 (item 84), .2957 (item 23). Item 5 was a cloze test grammar item, items 19, 21, 23, 33 were discrete point grammar items and items 84, 86 and 87 were sentential level fill in the blanks type vocabulary items. The easiest items had item difficulties of .9574 (item 43), .8831 (item 85), .8294 (item 51), .8212 (item 68), .8157 (item 26), .8088 (item 62). Item 26 was a discrete point grammar item, items 43, 51, 62 and 68 were sentence completion and reference type reading items, item 85 was a sentential level fill in the blanks type vocabulary item. Since there are not too many difficult and easy items and the mean of the total scores is 53.96, BUSPE 2001 can be considered to be at moderate difficulty.

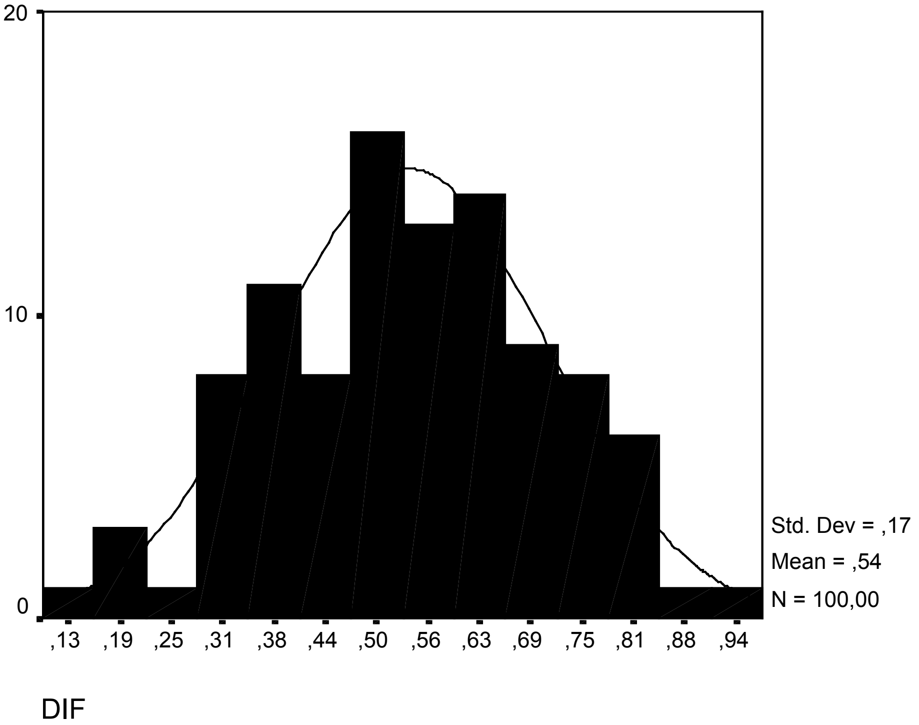

Figure 4.1.2 Frequency Distribution of Difficulty Indices of BUSPE 2001's Items

The discrimination indices of BUSPE 2001 ranged from -.0304 to .6423 with a mean of .3555 for the whole test. The means of item discrimination indices were .3238 for the grammar subtest, .4168 for the reading subtest, .2929 for the vocabulary subtest. The reading subtest was most powerful in discriminating among high and low performing students, whereas the vocabulary subtest was the least powerful. Since items with a discrimination index below .20 are considered as non or very poorly discriminating, fourteen items of BUSPE 2001 can be accepted as very poorly or non discriminating. Fourteen such items can be considered as tolerable in a 100-item test. These fourteen items had discrimination indices of -.0304 (item 19), .0258 (item 87), .0338 (item 94), .1185 (item 37), .1262 (item 23), .1363 (item 89), .1365 (item 61), .1374 (item 22), .1431 (item 39), .1434 (item 100), .1462 (item 33), .1556 (item 20), .1646 (item 97), .1742 (item 81). Items 19, 20, 22, 23 and 33 were discrete point grammar items. Items 37 and 39 were spot the mistake type grammar items, item 61 was a sentence completion type reading item, and items 81, 87, 89, 94, 97 and 100 were sentential level fill in the blanks type vocabulary items. The highest discrimination indices were .5006 (item 24), .5013 (item 71), .5069 (item 90), .5195 (item 69), .5196 (item 44), .5326 (item 70), .5343 (item 77), .5490 (item 42), .5534 (item 76), .5677 (item 17), .5763 (item 66), .6423 (item 12). Item 12 was a cloze test grammar item, items 17 and 24 were discrete point grammar items, items 42, 44, 66, 69, 70, 71, 76 and 77 were sentence completion and reference type reading items. Item 90 was a sentential level fill in the blanks type vocabulary item.

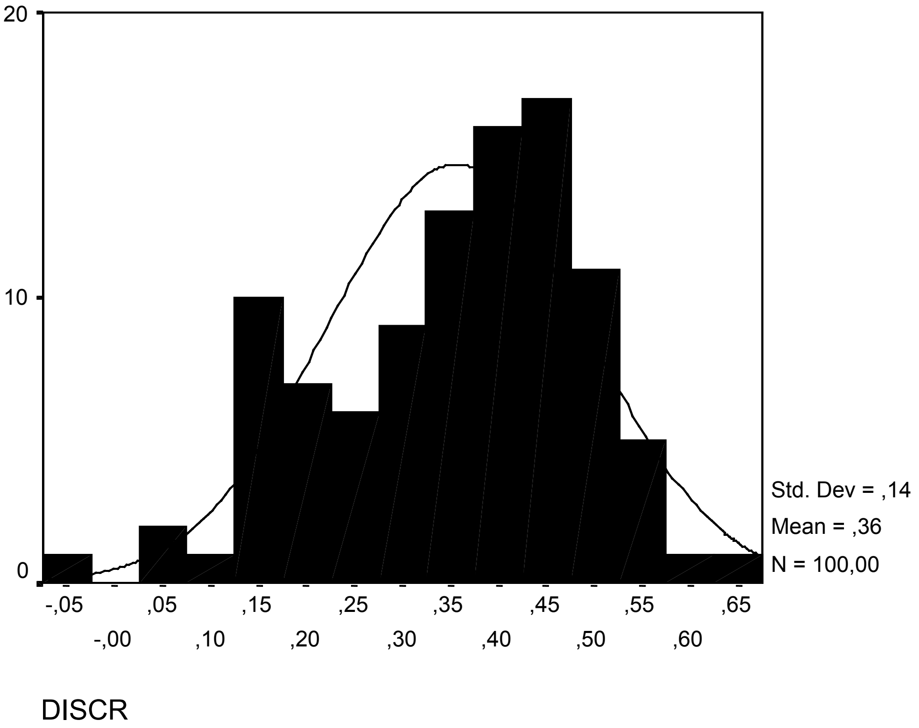

Figure 4.1.3 Frequency Distribution of Discrimination Indices of BUSPE 2001's Items

The coefficient alpha for the overall BUSPE 2001 was .9396. This alpha level is quite high and provides sufficient evidence to claim that BUSPE 2001 yields reliable scores. Moreover, if single items were deleted alpha did not differ much. It increased the most when items 19 (to .9405), 37 (to .9400), 87 (to. 9401), and 94 (to .9403) were deleted. These items were those which also had the lowest discrimination indices: -.304, .1185, .025, .0338.

4.2 Goodness of Fit Analyses

4.2.1 Checking Model Assumptions

The results of the principal component analysis indicated that BUSPE 2001 was unidimensional. There was a dominant factor with an eigenvalue of 16.337 accounting for a variance of 16.337 %. A second factor had an eigenvalue of 3.599 accounting for a variance of 3.599 % (Appendix 3). All factor were displayed in a scree plot. As seen in Figure 4.2.1.1 the first factor is about five times larger and there is a sharp fall from the first to the second eigenvalue supporting the unidimensonality assumption.

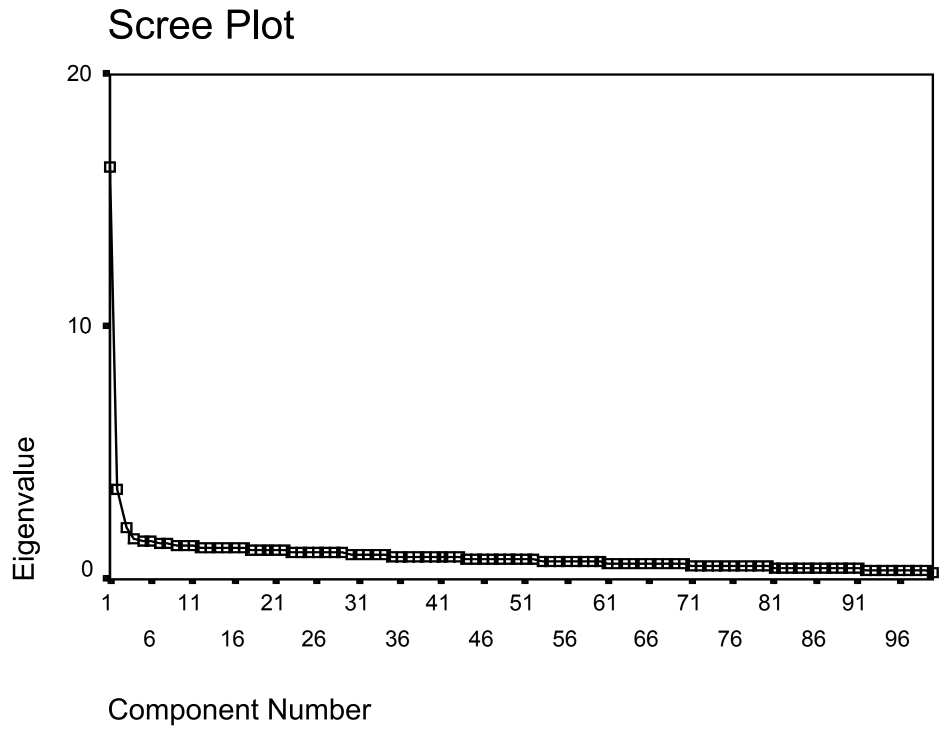

Figure 4.2.1.1 Scree Plot of Eigenvalues

The mean of inter-item correlation for all students was .1342, while it was .0113 for the group of students in the first quartile and .0202 for that of in the fourth quartile. The means of inter-item correlation for students in restricted range ability groups was lower than that for the whole group. Thus, it was found that the items of BUSPE 2001 were locally independent proving that unidimensionality entails local independence.

When figure 4.1.3, which is displaying the frequency distribution of the item discrimination indices of BUSPE 2001, is reexamined it can be seen that that the distribution was not homogeneous. The discrimination indices ranged from -.0304 to .6423 with a mean of .3555. The variance was .0184 and the standard deviation .1356. Hence, the data did not meet the assumption of equal discrimination indices of the one parameter IRT model.

The difficulty indices of the most difficult eight items were calculated for the low ability students who were in the first quartile in terms of total scores. The difficulty indices were .1094 for item 5, .3073 for item 19, .1042 for item 21, .2292 for item 23, .2135 for item 33, .1146 for item 84, .0078 for item 86 and .2031 for item 87. According to Table 4.2.1, the low ability students' performance on most difficulty items was worse in comparison with the whole group. However, they performed better on item 19, 33 and 87. These items could be misconstructed, since they were difficult, poorly discriminating and not loading on the dominant factor extracted by the principal component analysis. The low ability students performed worse on the other five difficult items and their difficulty indices for this group approached zero. Therefore, pseudo-chance factor could be taken into consideration only for some items.

Table 4.2.1.1 Performance of Low Ability Students on Difficult Items

ITEM	DIFF. WHOLE GROUP	DIFF. LOW GROUP	PERCENT OF INCORRECT WHOLE GROUP	PERCENT OF INCORRECT LOW GROUP
5	.2682	.1094	73.18	89.06
19	.2889	.3073	71.11	69.27
21	.2160	.1042	78.40	89.58
23	.2957	.2292	70.46	77.08
33	.1967	.2135	80.33	78.65
84	.2930	.1146	70.70	88.54
86	.1348	.0078	86.52	99.22
87	.2077	.2031	79.23	76.96

The variance of omitted items was found to be 8.0667, and the variance of incorrect responses was 43.9574. Their ratio was .23, which is close to zero indicating that BUSPE 2001 was not speeded.

4.2.2 Checking Model Predictions of Actual and Simulated Test Results

The overall chi square statistics was χ^2=.0000 significant at alpha level .05 for both the one parameter model and the two parameter model. 43 items did not fit both the one and two parameter model. At this stage, since the data seemed to fit neither of the IRT models, the data was reduced to 669 by excluding students with total scores below 30. This reduction was possible, because total scores below 30 are not very meaningful on a hundred-item multiple choice test, the items of which have only four choices and where there is no score reduction for wrong responses. The overall chi square statistics obtained with the new data for the one parameter model was χ^2=.0000 significant at alpha level .05. The number of items that were not fitting the one parameter model was 42 (Appendix 4). However, the overall chi square statistic became χ^2=.0911 insignificant at alpha level .05. The number of items not fitting the two parameter model decreased and became 8. These items were items 19 (χ^2=.0181), 33 (χ^2=.0002), 52 (χ^2=.0188), 57 (χ^2=.0242), 66 (χ^2=.0042), 75 (χ^2=.0221), 82 (χ^2=.0054) (Appendix 4). Thus, the two parameter model seems to fit the test data of BUSPE 2001 better than the one parameter as far as chi square statistics are considered.

The one parameter model provided information with minimum error approximately between ability levels –1.00 and +.05. Maximum information provided was approximately 12.0600 at ability level -1.9286. The two parameter model provided maximum information approximately between ability levels -.08 and .00 with minimum error. The two parameter model provided maximum information of approximately 30.2800 at ability level -2.0714. The two parameter model provided higher and more accurate information than the one parameter model (Appendix 5).

The ability estimates of the one and two parameter model were correlated with the total scores of BUSPE 2001 (Table 4.2.2.1). The ability estimates of the one parameter model had a correlation of .998 and those of the two parameter had a correlation of .974 with the total test scores. Both of these correlations were significant at alpha level .01 and they met the expectation that ability estimates of IRT models should have a strong relationship with actual test scores. These correlations are shown in scatter diagrams in Figure 4.2.2.1 and Figure 4.2.2.2.

Table 4.2.2.1 Correlation Between Ability Estimates of the 1P and 2P Models with the Total Test Scores

	ABILITY ESTIMATES (1 P. M.)	ABILITY ESTIMATES (2 P. M.)
TOTAL TEST SCORES	.998**	.974**

**Significant at α=.01

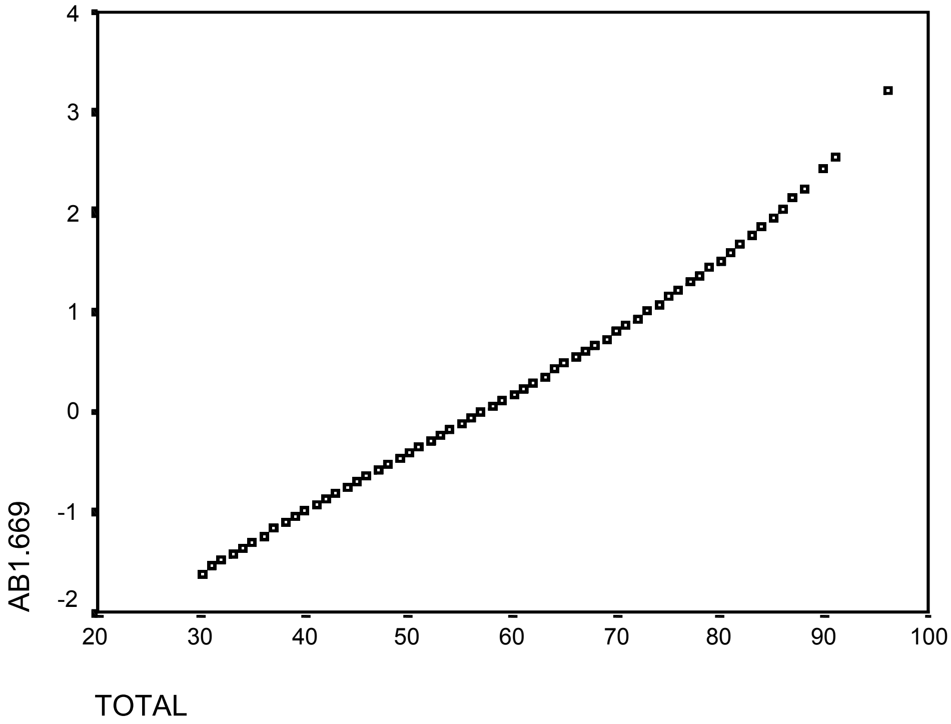

Figure 4.2.2.1 Scatter Diagram of Correlation Between Ability Estimates of the 1P Model and Total Test Scores

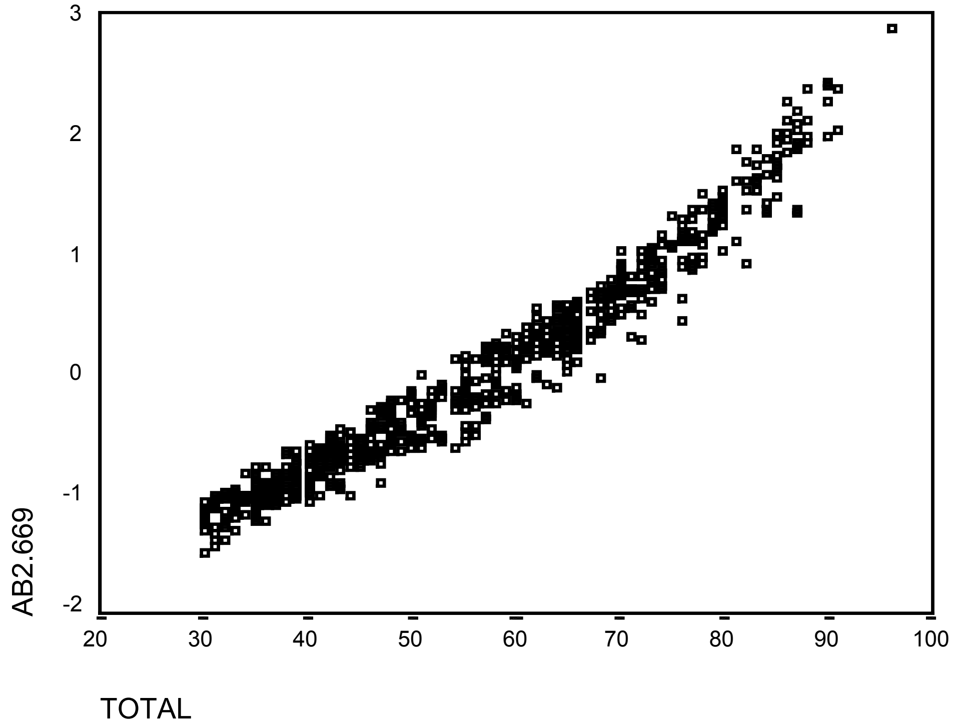

Figure 4.2.2.2 Scatter Diagram of Correlation Between Ability Estimates of the 2P Model and Total Test Scores

4.2.3 Checking Expected Model Features

The correlations between ability estimates obtained by easy and difficult items, and by odd and even items for both the one and two parameter IRT model are presented in Table 4.2.3.1. As indicated in Table 4.2.3.1 the correlation between odd and even abilities was .874 in the one parameter model, on the other hand it was .899 and higher in the two parameter model. However, both correlations were significant at alpha level .01 and quite high. These correlations are represented in scatter diagrams in Figures 4.2.3.1 and 4.2.3.2. The correlation between abilities obtained by easy and difficult items in the one parameter model was .835, while it was .867 in the two parameter model. These correlations are displayed in scatter diagrams in Figures 4.2.3.3 and 4.2.3.4. Both correlations were significant at alpha .01. Nevertheless, the correlation seems to be slightly higher in the two parameter model. Thus, although all correlations were high, the two parameter model yielded more invariant ability estimates across different sets of items.

Table 4.2.3.1 Correlations of Odd-Even, Easy-Difficult Abilities in the 1P and 2P Models

SUBTESTS	ONE PARAMETER IRT MODEL	TWO PARAMETER IRT MODEL
ODD-EVEN	.874**	.899**
EASY-DIFFICULT	.835**	.867**

**Significant at α=.01

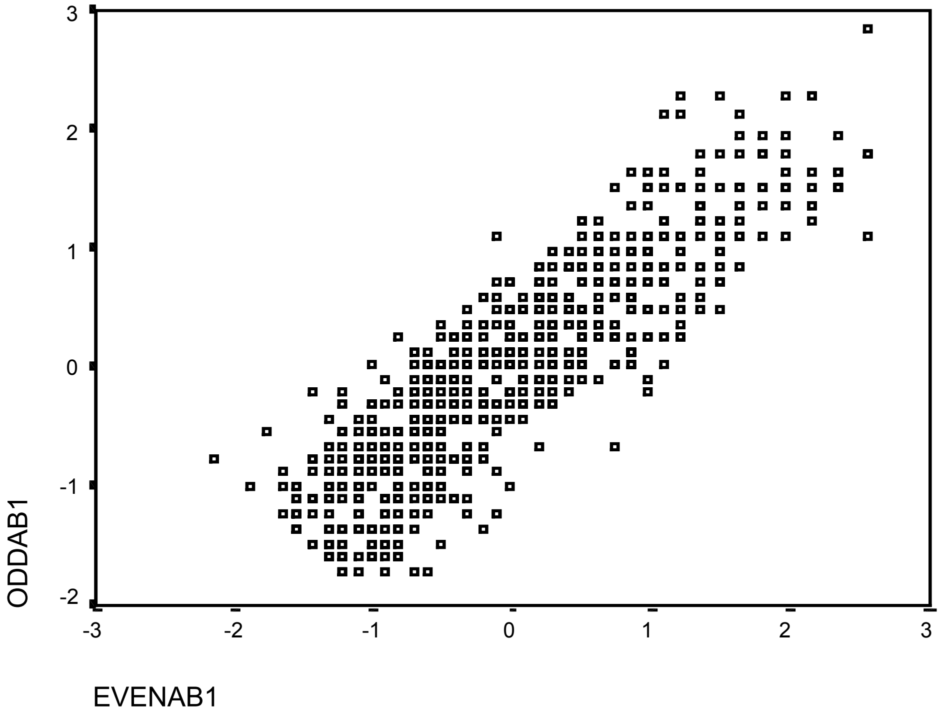

Figure 4.2.3.1 Scatter Diagram of Correlation Between Odd-Even Abilities in the 1P Model

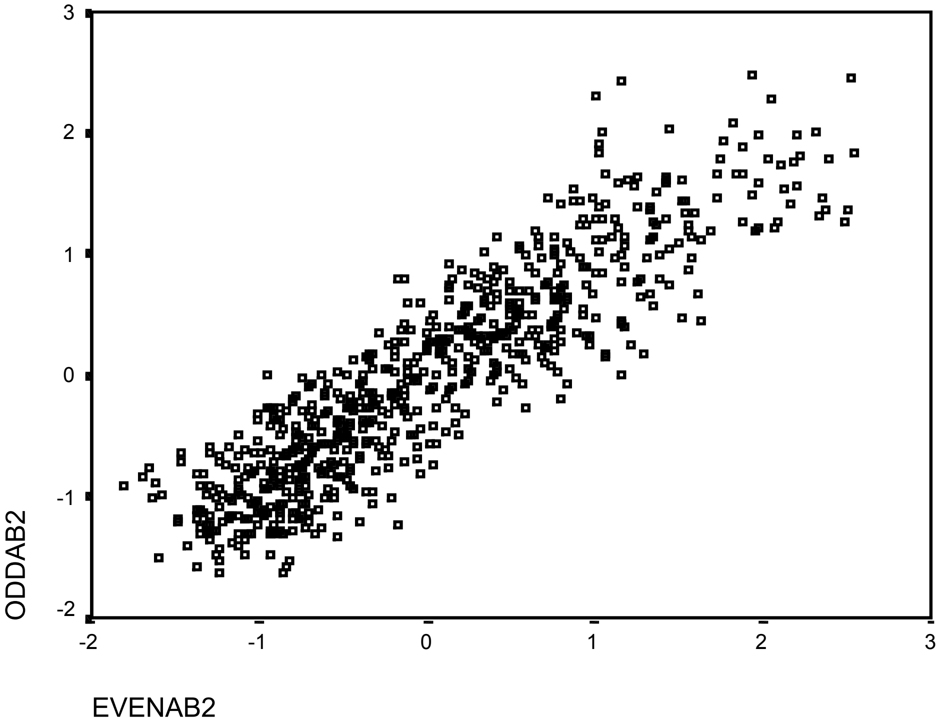

Figure 4.2.3.2 Scatter Diagram of Correlation Between Odd-Even Abilities in the 2P Model

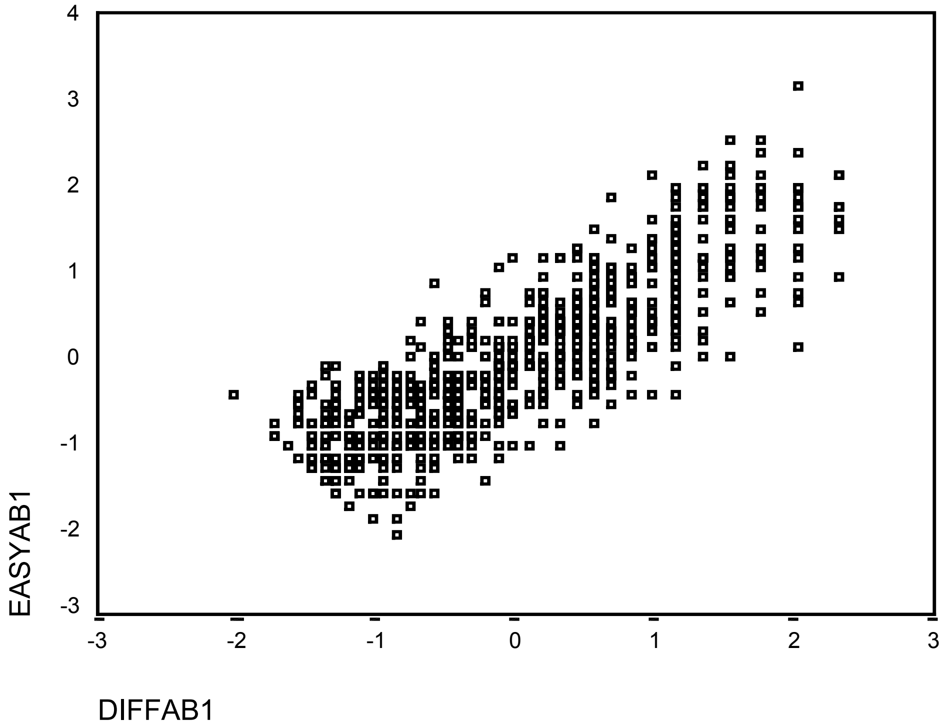

Figure 4.2.3.3 Scatter Diagram of Correlation Between Easy-Difficult Abilities in the 1P Model

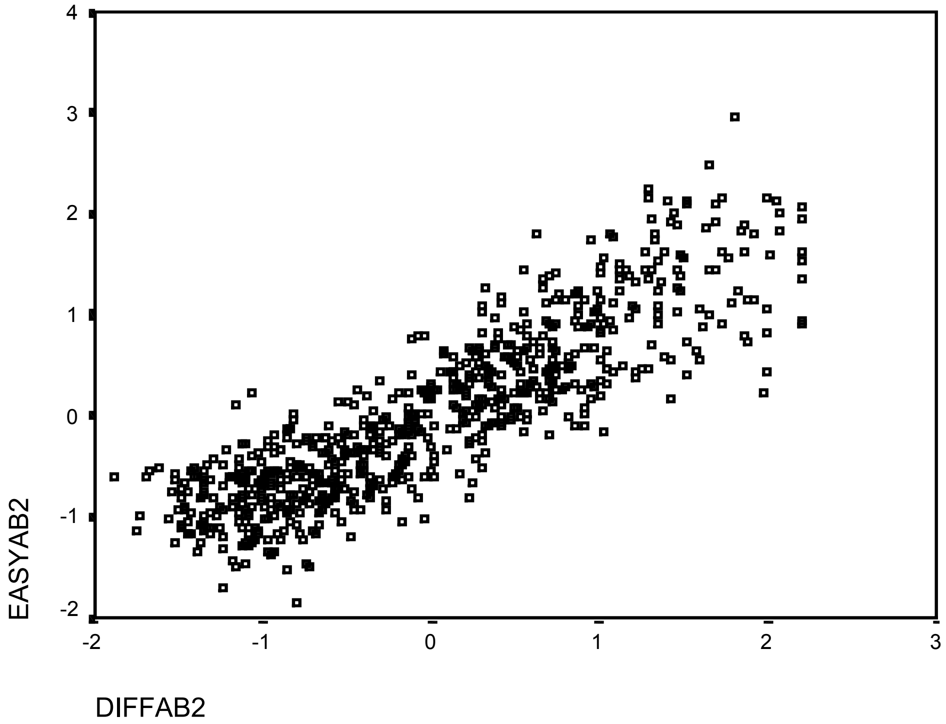

Figure 4.2.3.4 Scatter Diagram of Correlation Between Easy-Difficult Abilities in the 2P Model

The invariance of item parameters of the one and two parameter IRT models across groups of high and low performing students are presented in Table 4.2.3.2. In the one parameter IRT model the correlation between “b” parameter estimates obtained from high and low groups was .790 significant at alpha level .01 , in the two parameter IRT model the same correlation was .795 significant at alpha level .01. Figures 4.2.3.5 and 4.2.3.6 display these correlations in scatter diagrams. Both correlations were high, yet it was slightly higher in the two parameter IRT model.

The correlation between “a” parameter estimates of the two parameter IRT model across high and low performing groups was .569 significant at alpha level .01 (Figure 4.2.3.7).

Taking into consideration that high and low performing groups of students are challenging groups to analyze invariance of item parameters, it was observed that both IRT models provide invariant item parameters. However, the two parameter IRT model's "b" parameter estimates seem to be a little more invariant across different groups of students.

Table 4.2.3.2 Correlations of Item Parameter Estimates Across Different Samples of Students in the 1P and 2P Model

ITEM PARAMETERS & SAMPLES	ONE PARAMETER IRT MODEL	TWO PARAMETER IRT MODEL
"b" HIGH-LOW	.790**	.795**
"a" HIGH-LOW		.569**

**Significant at α=.01

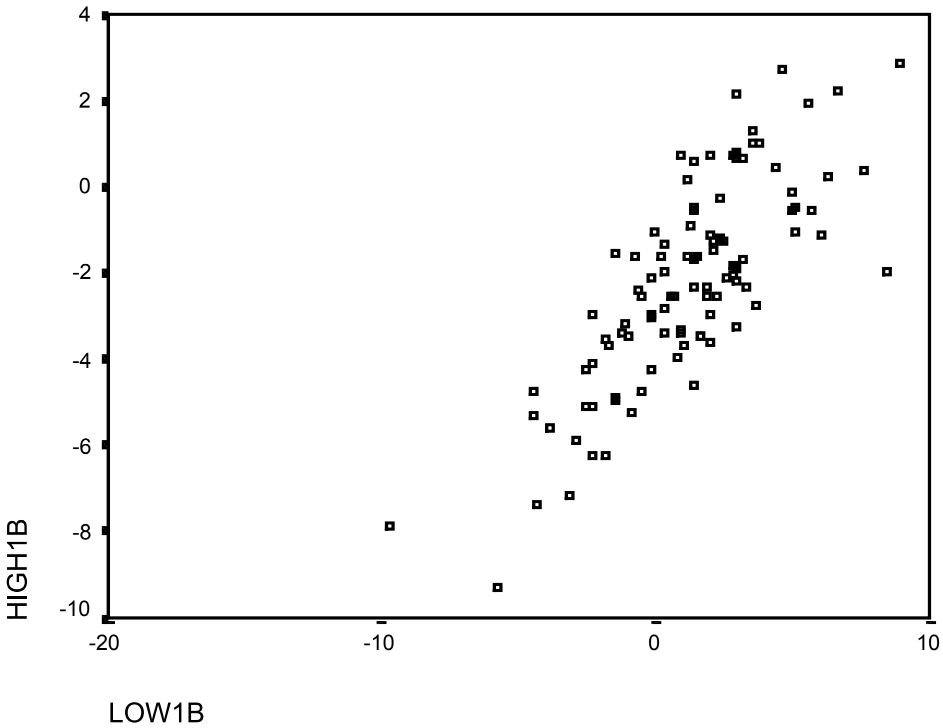

Figure 4.2.3.5 Scatter Diagram of Correlation between "b" Parameter Estimates of the 1P Model Obtained from High and Low Performing Students.

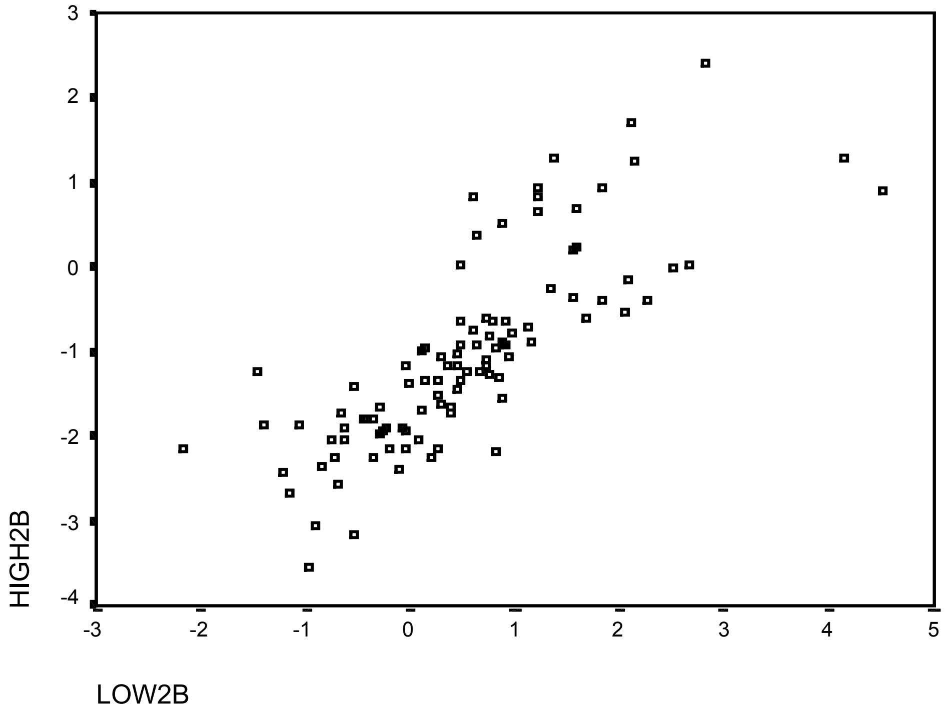

Figure 4.2.3.6 Scatter Diagram of Correlation between "b" Parameter Estimates of the 2P Model Obtained from High and Low Performing Students.

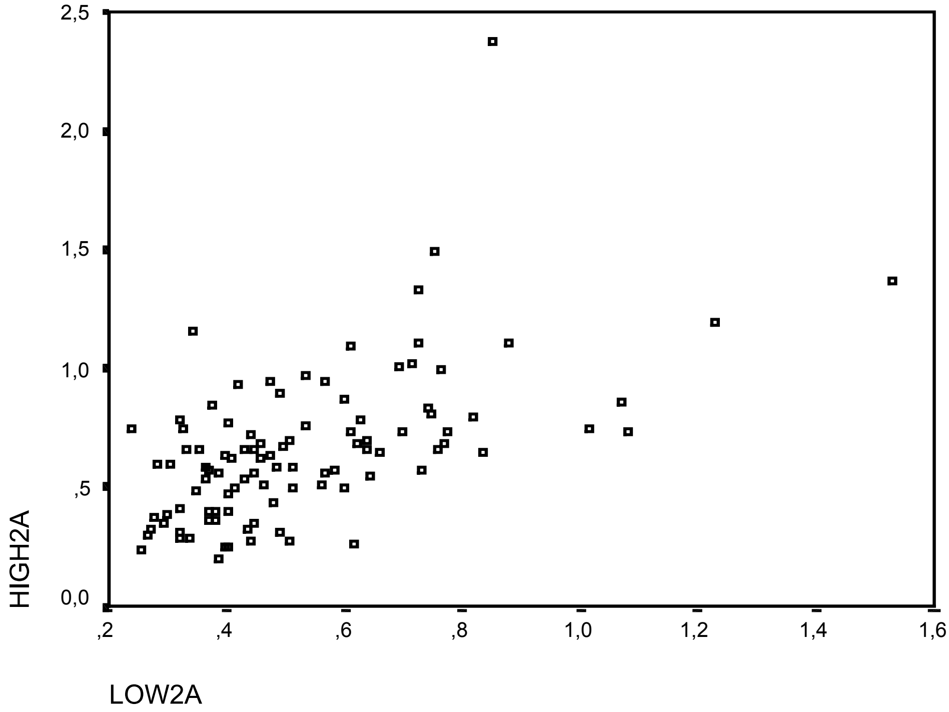

Figure 4.2.3.7 Scatter Diagram of Correlation between "a" Parameter Estimates of the 2P Model Obtained from High and Low Performing Students.

4.3 Predictive Validity Analyses

The correlations of BUSPE 2001's total scores and ability estimates of the two parameter IRT model with the DEC 1-2 grades are shown in Table 4.3.1. The total scores had a correlation of .693 with DEC 1 grades and .558 with DEC 2 grades, which are represented in scatter diagrams in Figure 4.3.1 and 4.3.2. On the other hand the ability estimates of the two parameter IRT model yielded correlations of .716 and .603 with DEC 1 and DEC 2 grades respectively. The scatter diagram of these correlations are displayed in Figure 4.3.3 and 4.3.4. All correlations were significant at alpha level .01. The ability estimates had higher correlations with DEC 1-2 grades than the total scores did. Both the ability estimates and the total scores had higher correlation with DEC 1 grades.

Table 4.1 Correlations of BUSPE 2001's Total Scores and 2P Abilities with DEC 1-2 Grades

	DEC 1	DEC 2
TOTAL SCORES	.693**	.558**
2P ABILITIES	.716**	.603**

**Significant at α=.01

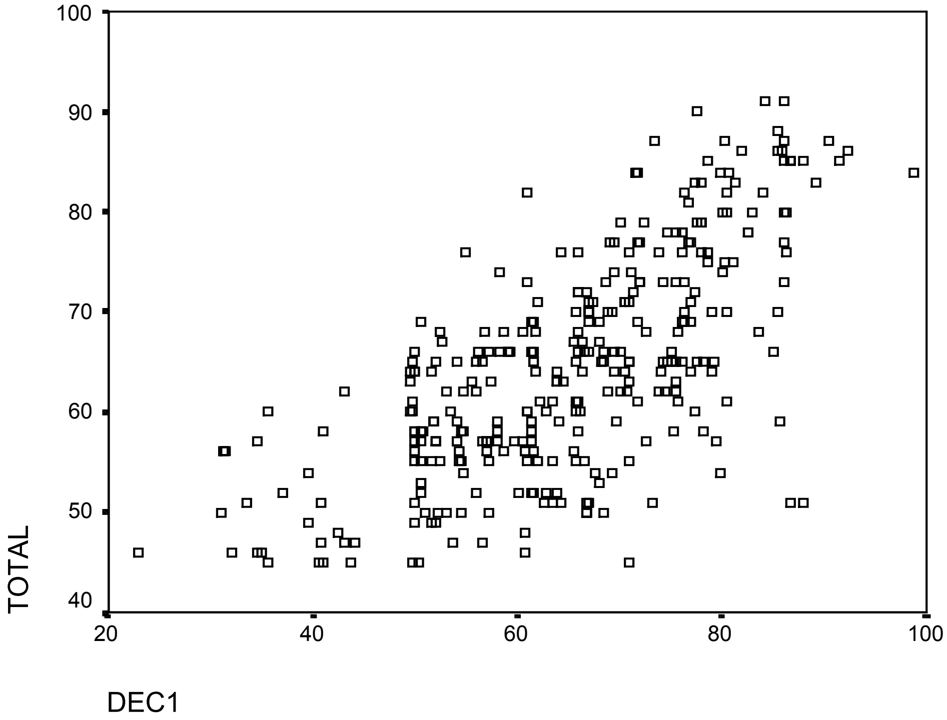

Figure 4.1 Scatter Diagram of Correlation Between Total Scores of BUSPE 2001 and Dec 1 grades

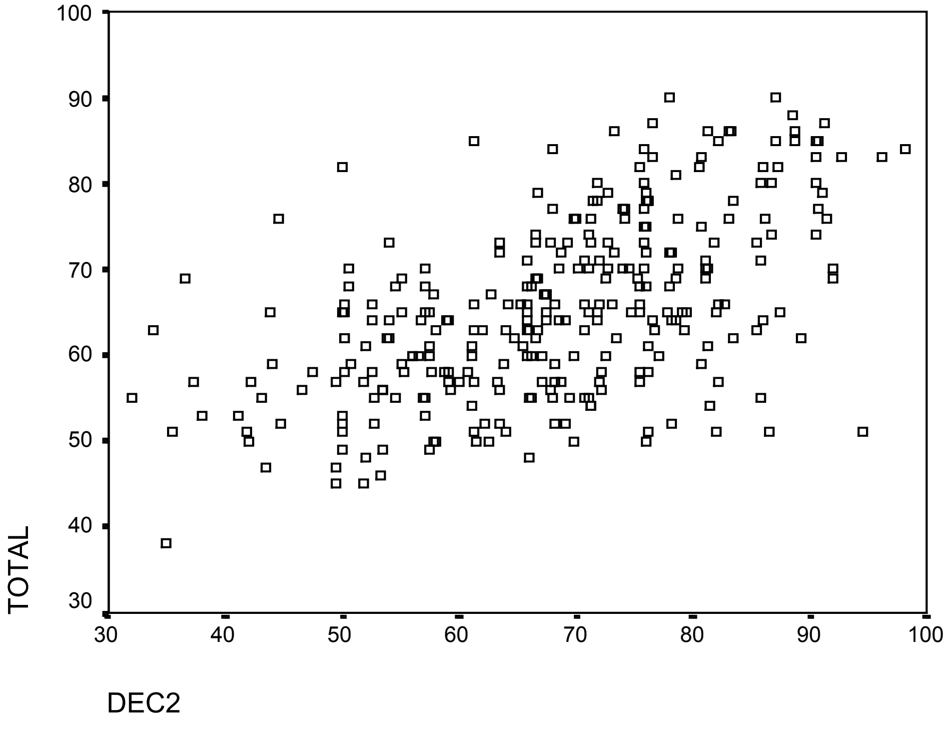

Figure 4.2 Scatter Diagram of Correlation Between Total Scores of BUSPE 2001 and Dec 2 grades

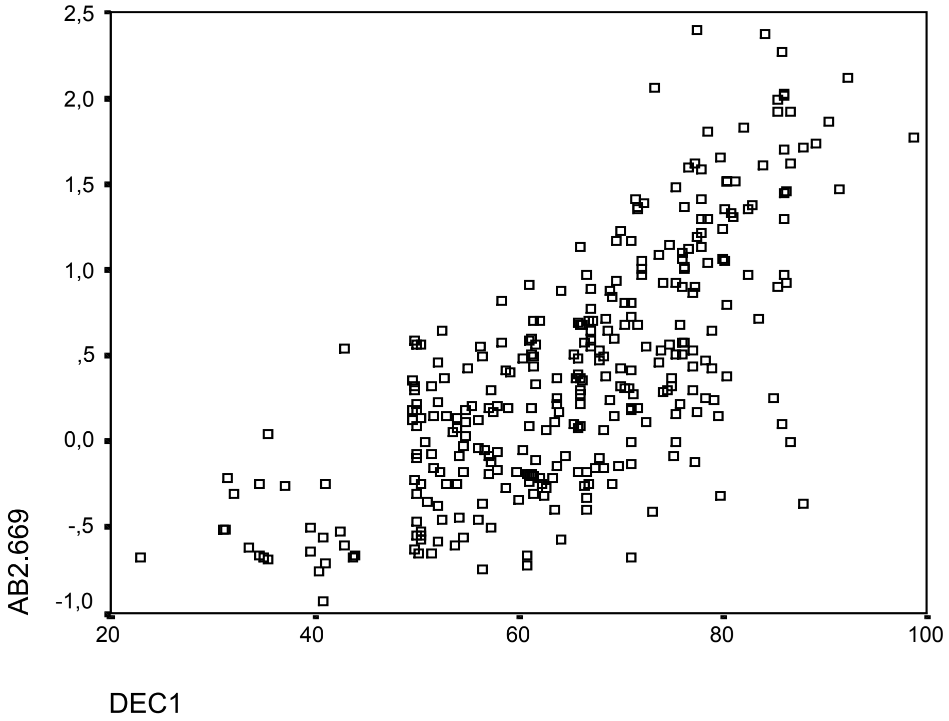

Figure 4.3 Scatter Diagram of Correlation between 2P Abilities and Dec 1 Grades

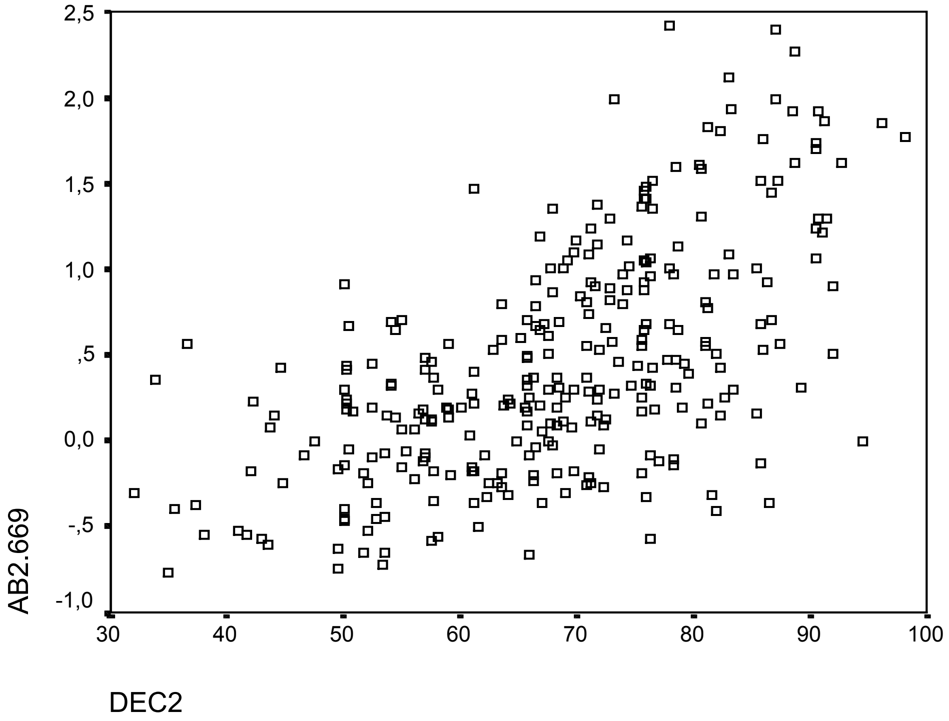

Figure 4.4 Scatter Diagram of Correlation between 2P Abilities and Dec 2 Grades

After checking the information indices and ICCs of each item, first fifty items providing highest information were detected. The information values of all items are presented in Table 4.2 and the ICCs of item 12 and item 94 are displayed in Appendix 1 as examples for good & informative, and poor and uninformative items. When the ICC of item 12 was examined it was observed that it had a very steep slope which meant that it was highly successful in discriminating among different ability students, whereas that of item 94 was not steep at all and had no function in discriminating among different ability students. Item 12 provided an information of 2.2462 while item 94 provided an information of .0292.

Table 4.2 Table of Item Information Values in the 2P Model

ITEM	INFO.	ITEM	INFO.	ITEM	INFO.	ITEM	INFO.
12	2,2462	62	,5019	83	,2634	33	,1040
43	2,1519	14	,4910	75	,2482	1	,0935
66	1,5059	5	,4889	74	,2404	7	,0870
85	1,3263	8	,4633	93	,2360	82	,0868
72	1,1653	4	,4586	10	,2304	27	,0761
79	1,0576	78	,4548	34	,2277	38	,0644
42	1,0166	13	,4416	35	,2233	40	,0635
48	,8498	56	,4295	64	,2226	73	,0558
77	,8269	47	,4275	2	,2194	99	,0520
51	,8186	32	,4253	6	,2100	39	,0493

ITEM	INFO.	ITEM	INFO.	ITEM	INFO.	ITEM	INFO.
17	,8039	53	,4159	96	,2074	21	,0468
76	,7935	11	,4018	58	,2070	81	,0421
70	,7906	55	,3976	30	,2069	61	,0418
59	,7861	54	,3657	46	,1922	31	,0401
44	,7722	84	,3622	57	,1914	25	,0388
95	,7465	65	,3566	88	,1895	100	,0346
69	,7228	15	,3471	60	,1816	20	,0336
71	,6797	45	,3462	86	,1624	37	,0292
24	,6728	41	,3289	26	,1601	22	,0288
67	,6493	68	,3135	49	,1562	89	,0251
9	,6324	63	,3038	18	,1180	97	,0245
90	,6118	36	,3007	98	,1143	23	,0237
29	,6027	3	,2767	16	,1095	19	,0207
80	,5770	92	,2739	91	,1068	87	,0169
50	,5609	52	,2665	28	,1047	94	,0102

Fifteen of the forty grammar items (% 37.5), thirty of the forty reading items (% 75) and ten of the twenty vocabulary items (% 50) were selected. Next, thirty items providing highest information were selected. This new set of items consisted of nine grammar items (%22.5), eighteen reading items (%45) and three vocabulary items (% 15). The reading items seemed to provide higher information than vocabulary and grammar items. This could be explained by the stronger realtionship between success in reading and overall proficiency which is necesarry to succeed in DEC. The reliability coefficient was calculated as .94 for the set of fifty items and .92 for the set of thirty items.

When the test information functions were observed (Appendix 5) it was found out that in the two parameter IRT model the set of fifty items providing highest information yielded information with minimum error approximately between ability levels –1.2 and +.03. Maximum information provided was approximately 16.96 at ability level -2.0000. On the other hand, the set of thirty items yielding highest information provided maximum information approximately between ability levels -1.00 and +.02 with minimum error. Maximum information provided was approximately 12.69 at ability level -2.000.

The total scores of the fifty items had a correlations of .680 and .572 with DEC 1-2 scores. The ability estimates obtained by these fifty items in the two parameter IRT model had correlations of .697 and .594 with DEC 1-2 scores. The total scores of thirty items providing highest information yielded correlations of .664 and .544 with DEC 1-2 scores. The ability estimates obtained by these thirty items in the two parameter IRT model had correlations of .679 and .569 with DEC 1-2 scores. It was observed that ability estimates obtained by both these fifty and thirty items had higher correlations with DEC 1-2 scores than their total and did. These ability estimates' correlations with DEC 1-2 scores were even higher that of BUSPE 2001's total scores. All of these correlations were significant at alpha level .01. The correlations are shown in Table 4.3 and scatter diagrams regarding them are displayed in figures 4.5, 4.5, 4.6, 4.7, and 4.8.

Table 4.3 Correlations of Total Scores and 2P Abilities of High Information Items with DEC 1-2 Grades

	DEC 1	DEC 2
TOTAL SCORES OF HIGH INF. 50 ITEMS	.680**	.572**
2P ABILITIES OF HIGH INF. 50 ITEMS	.697**	.594**
TOTAL SCORES OF HIGH INF. 30 ITEMS	.664**	.544**
2P ABILITIES OF HIGH INF. 30 ITEMS	.679**	.569**

**Significant at α=.01

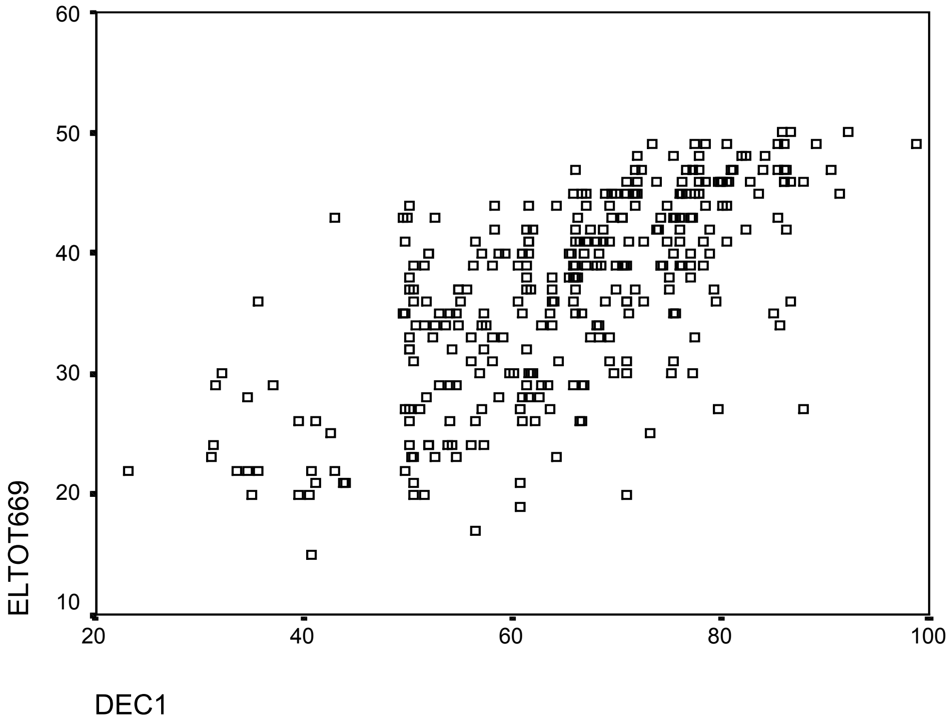

Figure 4.5 Scatter Diagram of Correlation Between Total Scores of 50 High Information Items and Dec 1 Grades

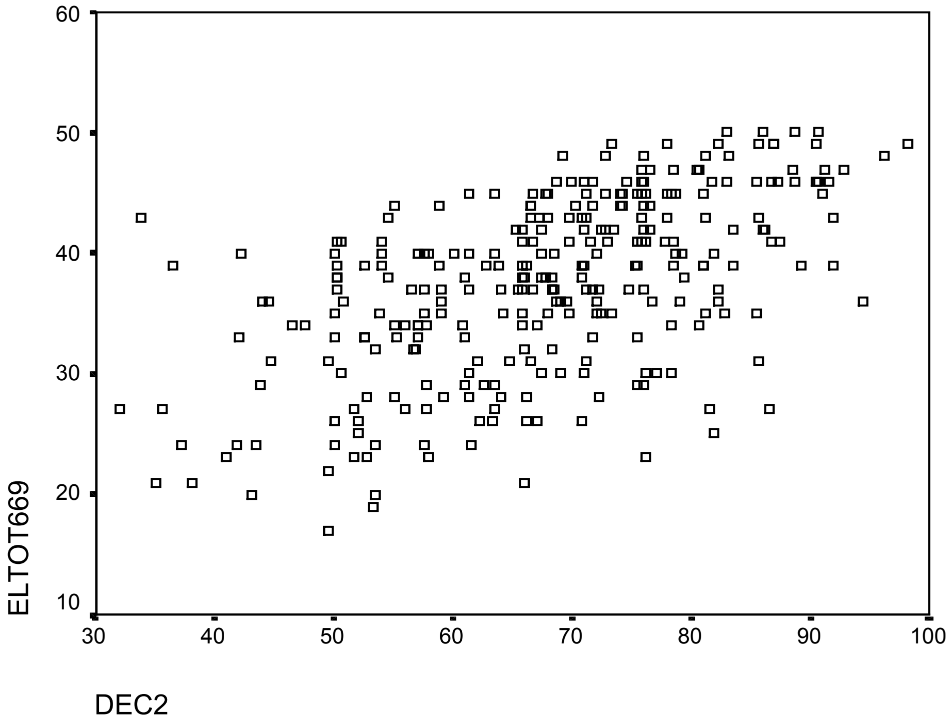

Figure 4.6 Scatter Diagram of Correlation Between Total Scores of 50 High Information Items and Dec 2 Grades

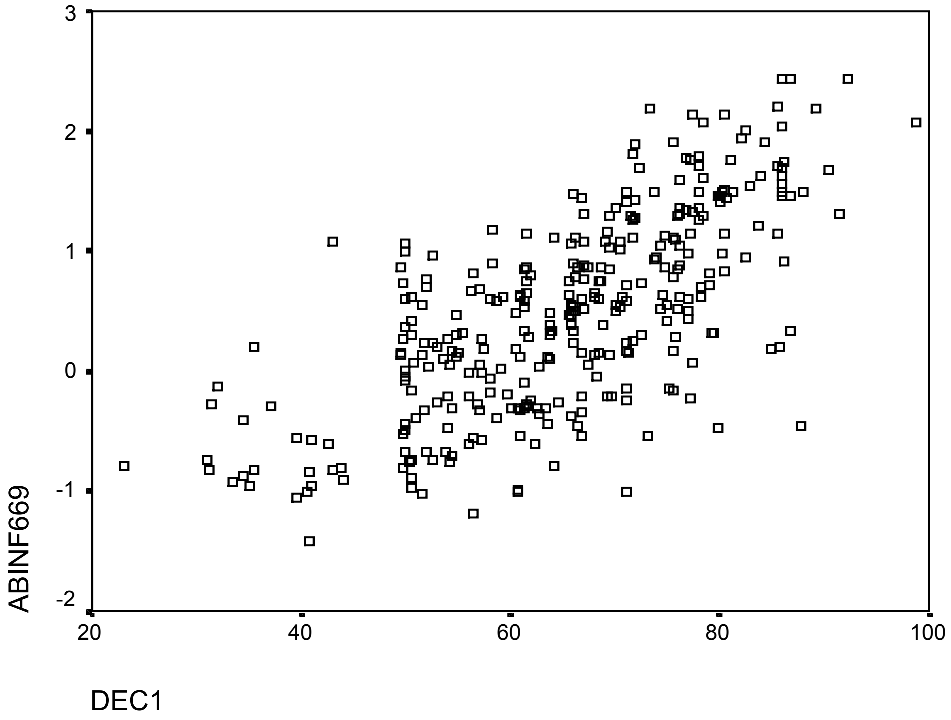

Figure 4.7 Scatter Diagram of Correlation Between 2P Abilities of 50 High Information Items and Dec 1 Grades

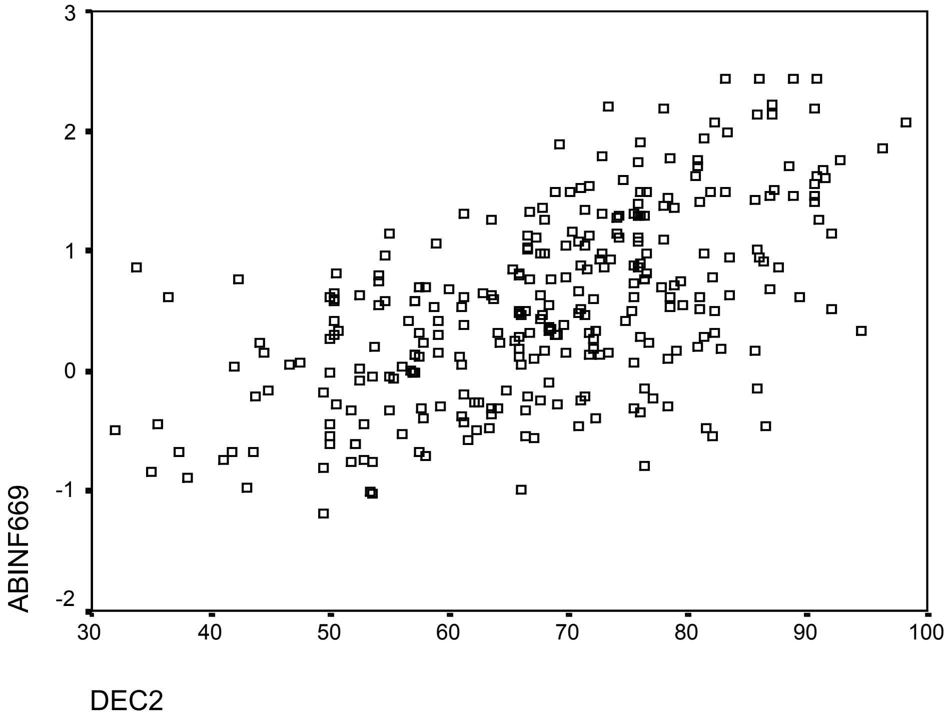

Figure 4.8 Scatter Diagram of Correlation Between 2P Abilities of 50 High Information Items and Dec 2 Grades

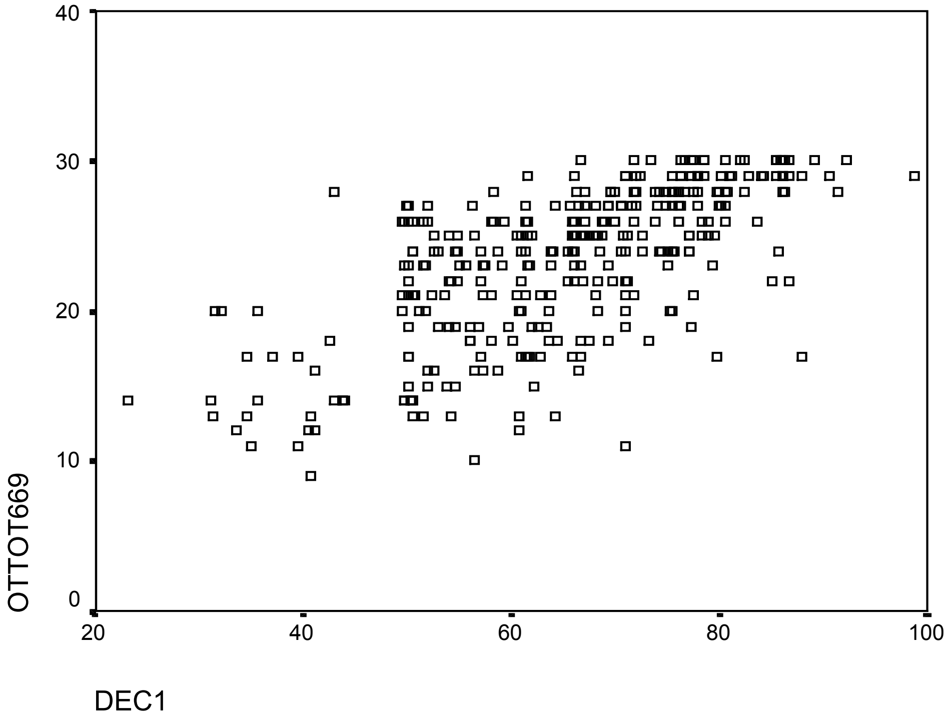

Figure 4.9 Scatter Diagram of Correlation Between Total Scores of 30 High Information Items and Dec 1 Grades

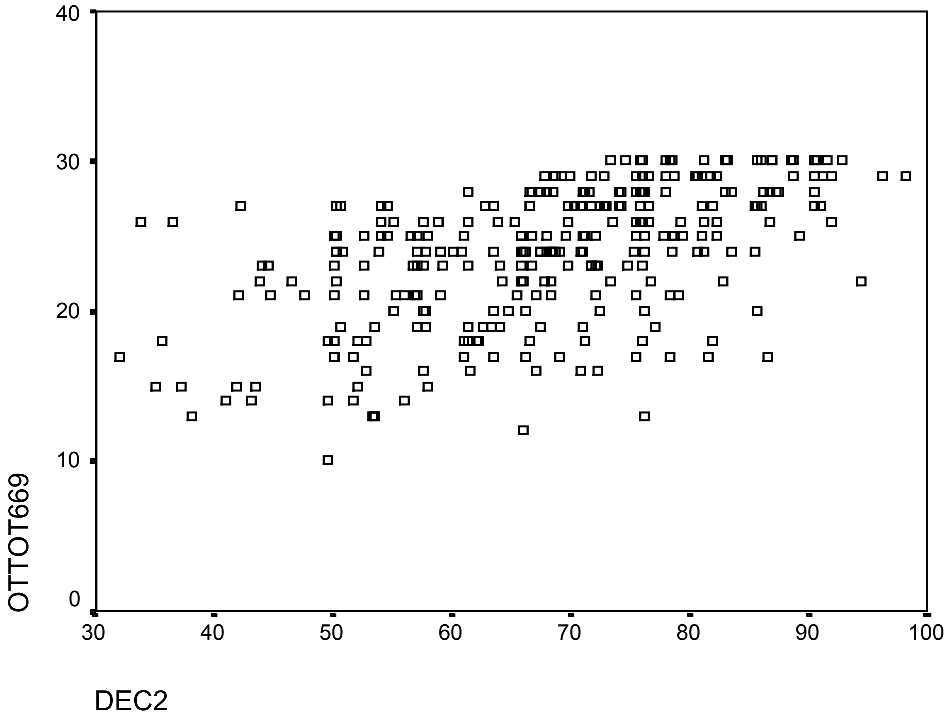

Figure 4.10 Scatter Diagram of Correlation Between Total Scores of 30 High Information Items and Dec 2 Grades

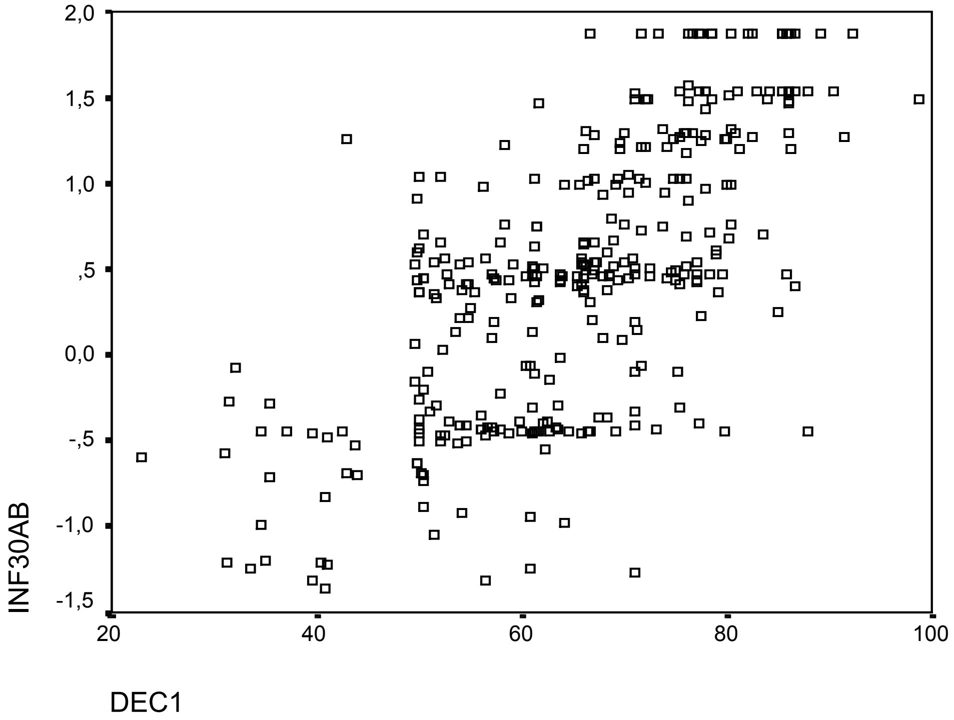

Figure 4.11 Scatter Diagram of Correlation Between 2P Abilities of 30 High Information Items and Dec 1 Grades

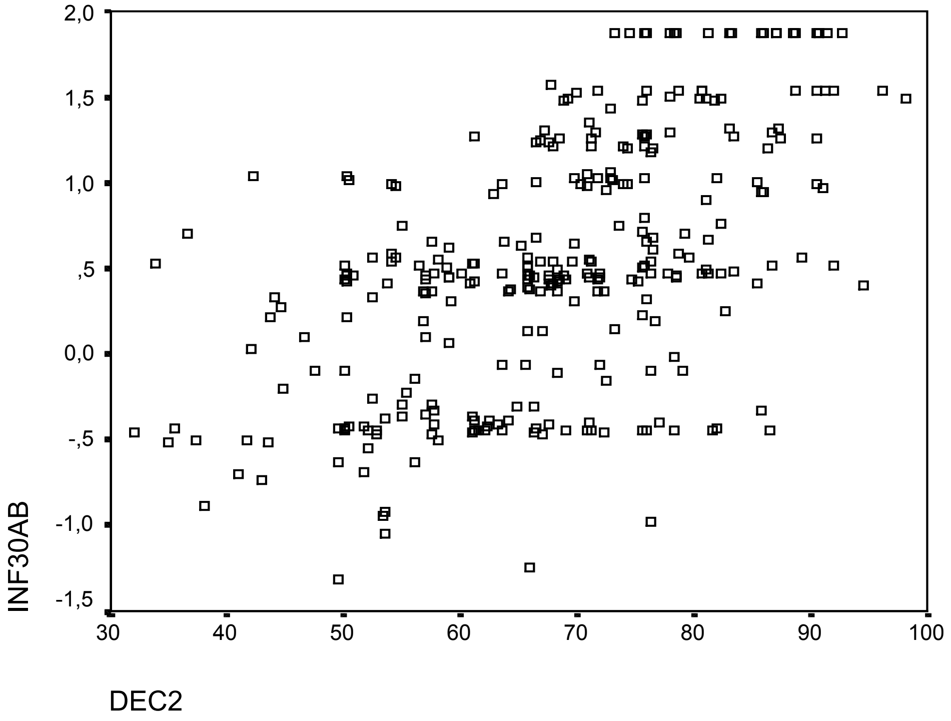

Figure 4.12 Scatter Diagram of Correlation Between 2P Abilities of 30 High Information Items and Dec 2 Grades

The correlations of the grammar, reading and vocabulary subtests' total scores and ability estimates with DEC 1-2 scores are presented in Table 4.4. The total scores of the grammar, reading and vocabulary subtest had correlations of .526, .653, .292 with DEC 1 scores and .441, .512, .228 with DEC 2 scores respectively. The ability estimates obtained by the grammar, reading and grammar subtests in the two parameter IRT model had correlations of .601, .676, .397 with DEC 1 scores and .528, .538, .316 with DEC 2 scores. Both the ability and total scores of the reading subtest had higher correlations with DEC 1-2 scores compared to the grammar and vocabulary subtests. Moreover, the subtests' ability estimates yielded higher correlations with DEC 1-2 scores than their total scores did. All correlations were significant at alpha level .01. Scatter diagrams of these correlations are displayed in Figures 4.13, 4.14, 4.15, 4.16, 4.17, 4.18, 4.19, 4.20, 4.21, 4.22, 4.23 and 4.24.

Table 4.4 Correlations of the Grammar, Reading and Vocabulary Subtests' Total Scores and 2P Abilities with DEC 1-2 Scores

	DEC 1	DEC 2
TOTAL SCORES OF GRAMMAR SUBTEST	.526**	.441**
TOTAL SCORES OF READING SUBTEST	.653**	.512**
TOTAL SCORES OF VOCAB. SUBTEST	.292**	.228**

	DEC 1	DEC 2
2P ABILITIES OF GRAMMAR SUBTEST	.601**	.528**
2P ABILITIES OF READING SUBTEST	.676**	.538**
2P ABILITIES OF VOCAB. SUBTEST	.397**	.316**

**Significant at α=.01

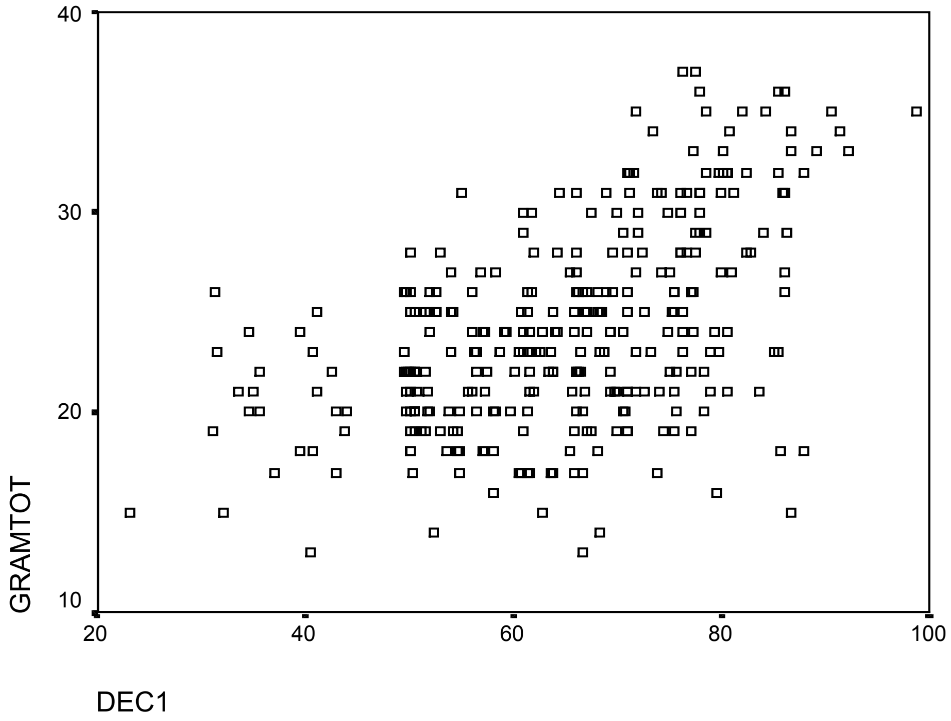

Figure 4.13 Scatter Diagram of Correlation between the Grammar Subtest's Total Scores and DEC 1 Scores

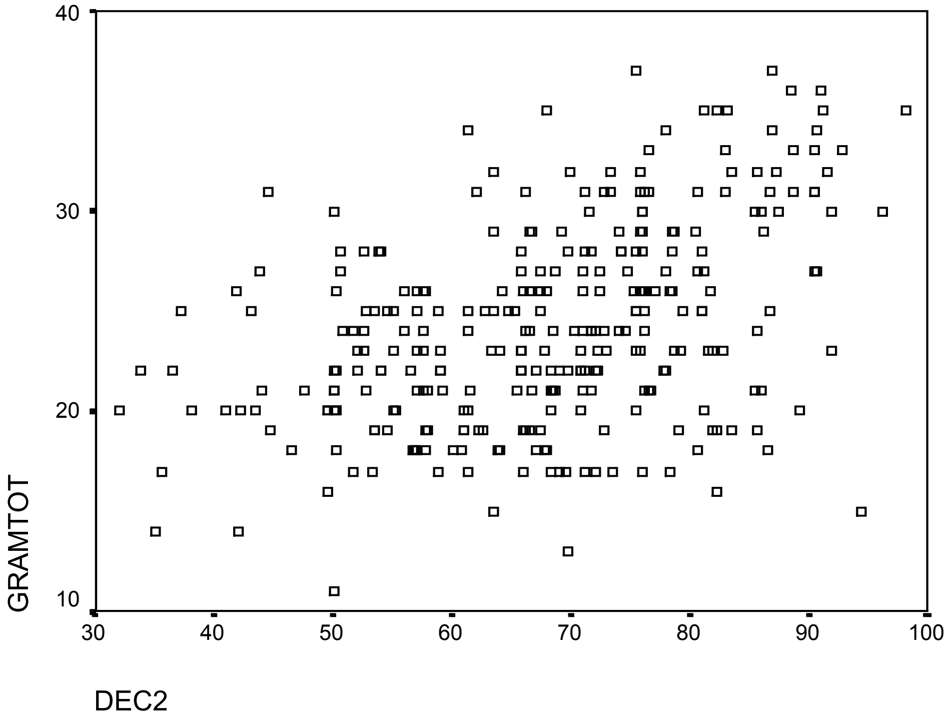

Figure 4.14 Scatter Diagram of Correlation between the Grammar Subtest's Total Scores and DEC 2 Scores

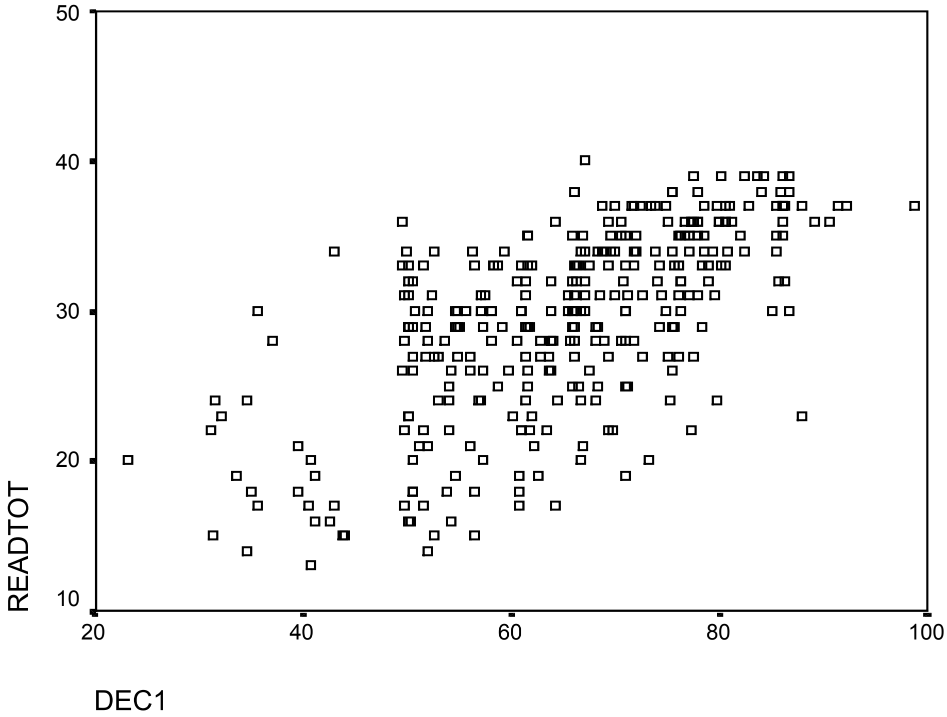

Figure 4.15 Scatter Diagram of Correlation between the Reading Subtest's Total Scores and DEC 1 Scores

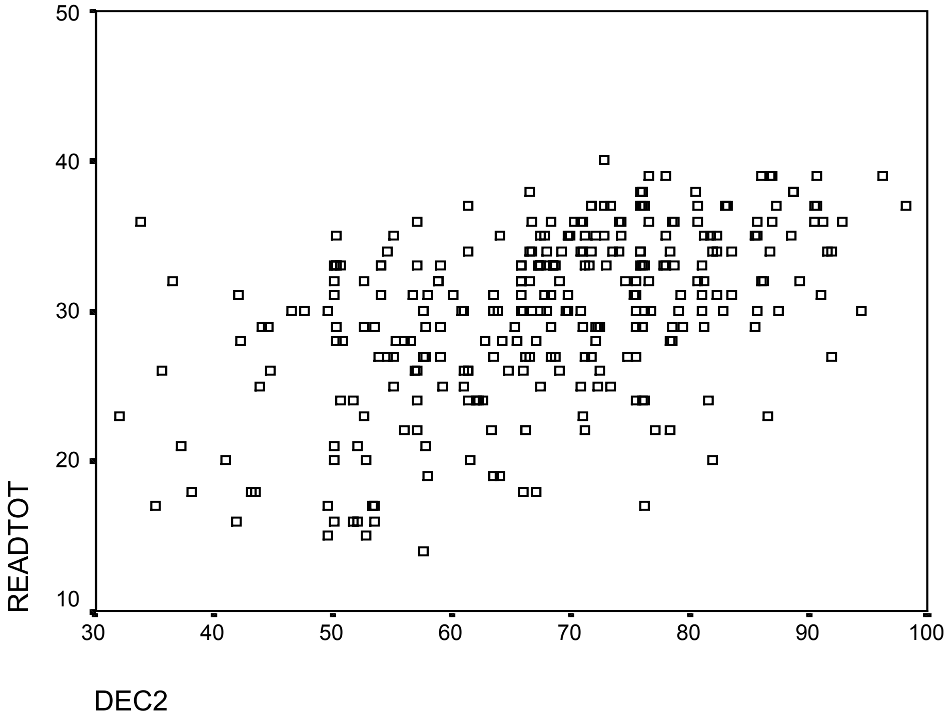

Figure 4.16 Scatter Diagram of Correlation between the Reading Subtest's Total Scores and DEC 2

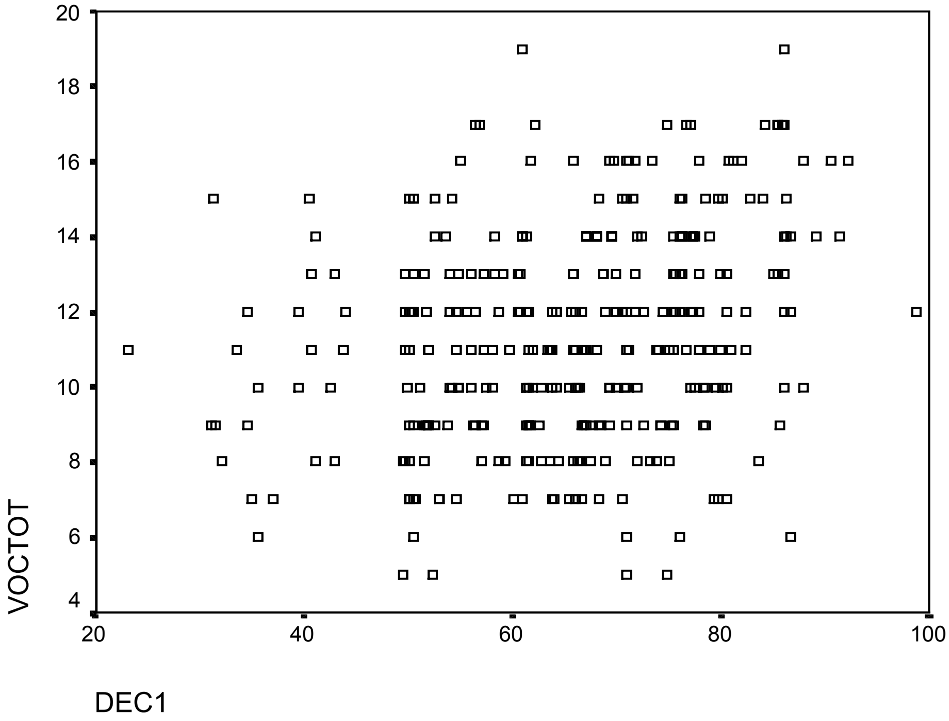

Figure 4.17 Scatter Diagram of Correlation between the Vocabulary Subtest's Total Scores and DEC 1 Scores

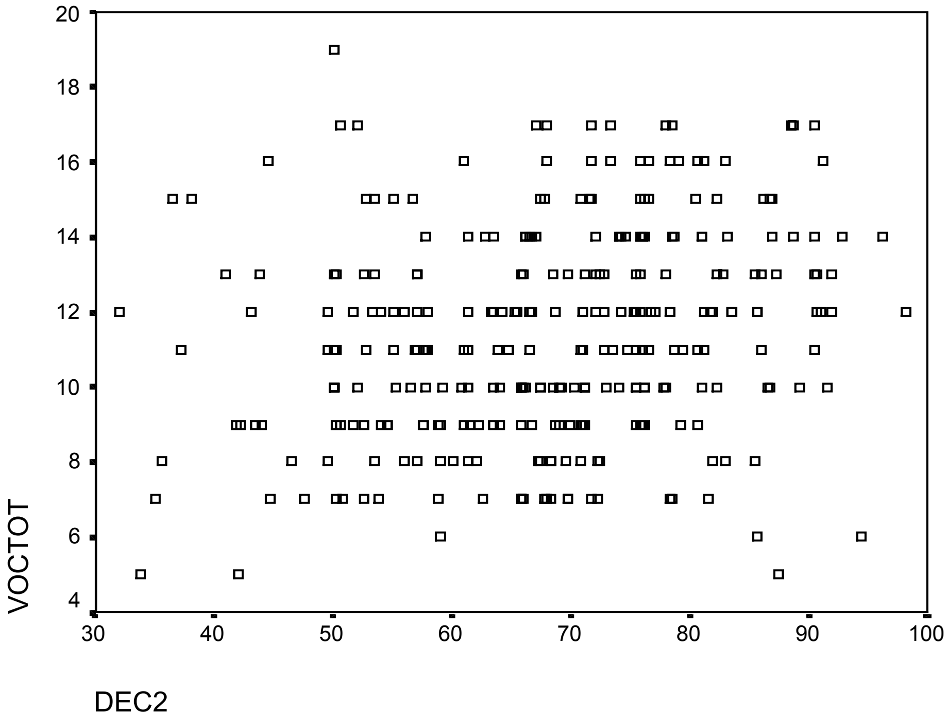

Figure 4.18 Scatter Diagram of Correlation between the Vocabulary Subtest's Total Scores and DEC 2 Scores

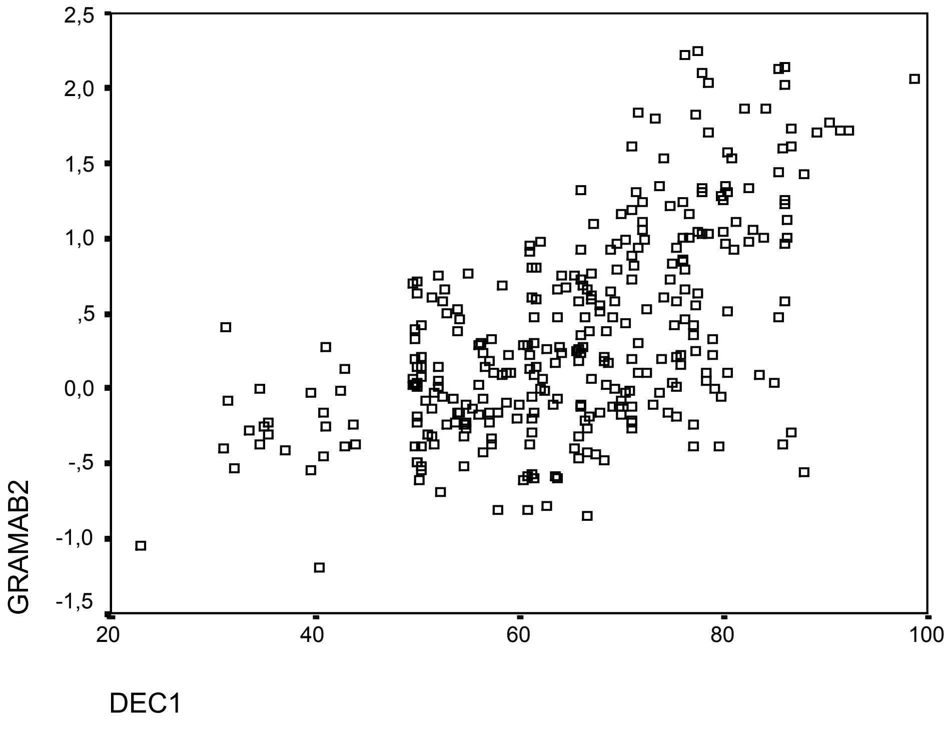

Figure 4.19 Scatter Diagram of Correlation between 2P Abilities of the Grammar Subtest and DEC 1 Scores

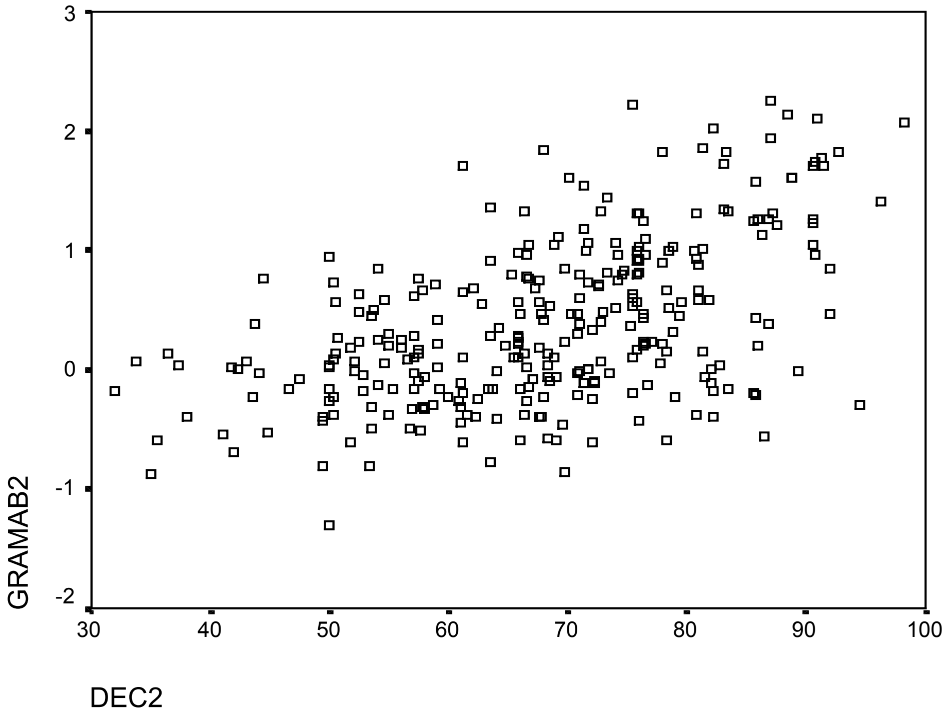

Figure 4.20 Scatter Diagram of Correlation between 2P Abilities of the Grammar Subtest and DEC 2 Scores

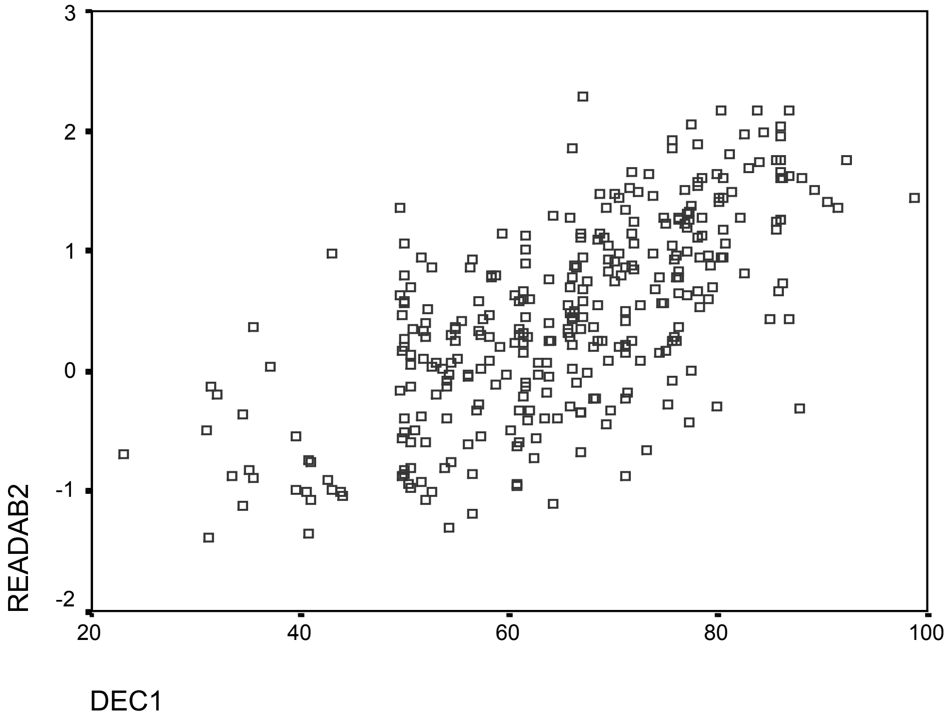

Figure 4.21 Scatter Diagram of Correlation between 2P Abilities of the Reading Subtest and DEC 1 Scores

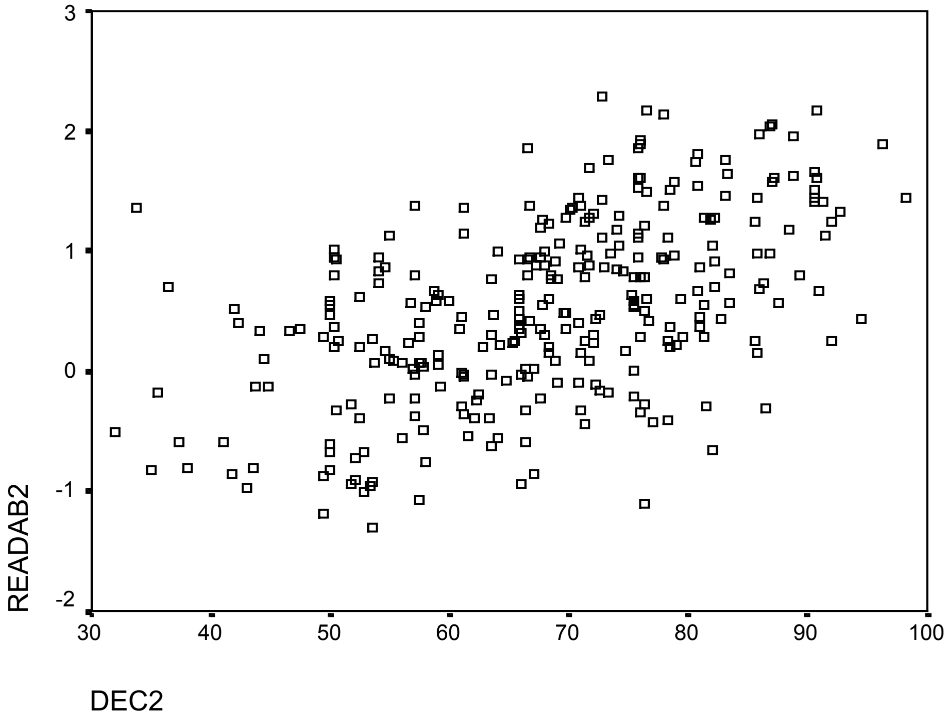

Figure 4.22 Scatter Diagram of Correlation between 2P Abilities of the Reading Subtest and DEC 2 Scores

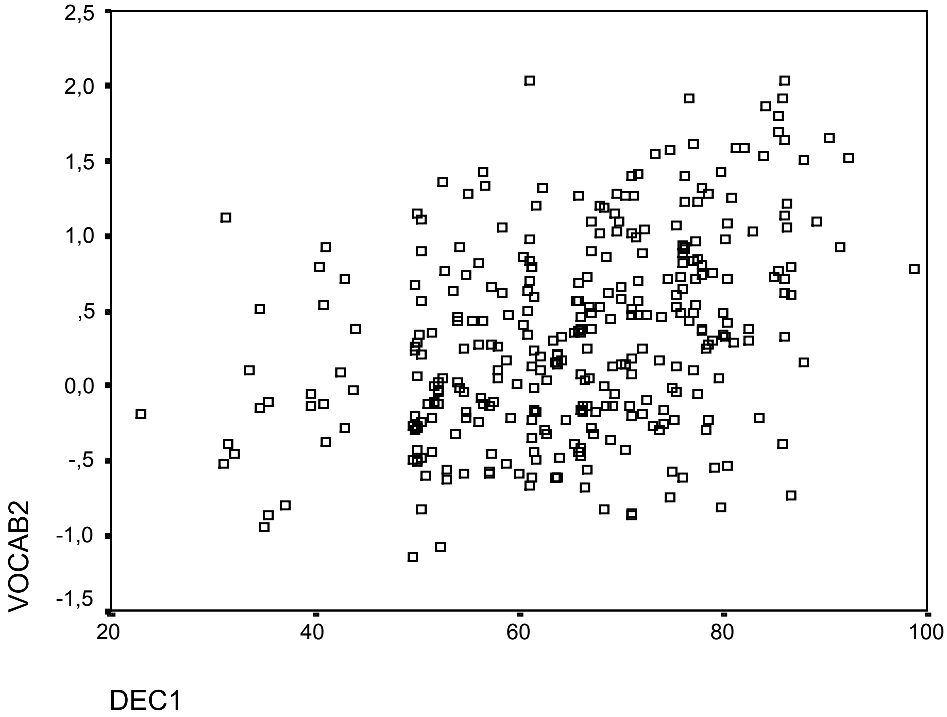

Figure 4.23 Scatter Diagram of Correlation between 2P Abilities of the Vocabulary Subtest and DEC 1 Scores

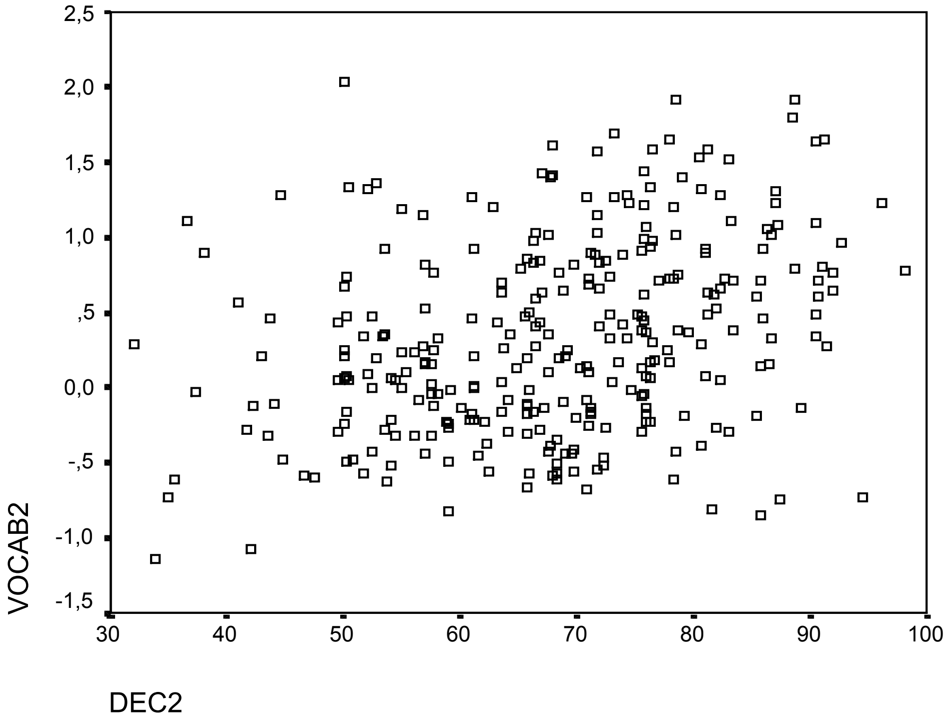

Figure 4.24 Scatter Diagram of Correlation between 2P Abilities of the Vocabulary Subtest and DEC 2 Scores

CHAPTER V

CONCLUSIONS & IMPLICATIONS

In this final chapter the discussions, the conclusions and implications of the study are presented.

5.1 Discussion

According to the descriptive statistics and the results of the classical item analysis, it can be concluded that BUSPE 2001 was moderately difficult for the students. BUSPE 2001 included both difficult and easy, both highly discriminating and poorly discriminating items.

The vocabulary subtest was found the most difficult by the students, the grammar subtest was less difficult, and the reading subtest was the easiest. On the contrary, the reading subtest was the most highly discriminating, the grammar subtest had less discriminative power while the vocabulary subtest was least discriminating.

BUSPE 2001 was proven to be highly reliable, because the coefficient alpha calculated for it was .94 where Alderson, Clapham, and Wall (1995) claim that a well-constructed and objective test of 100 multiple choice grammar items, which has been pretested on students with a wide range of language ability, might have a reliability index of .95. The fact that BUSPE 2001 was not pretested and that it did not include only grammar items makes it very successful in this sense. A contribution to its high reliability might have been from its unidimensionality which Alderson, Clapham, and Wall (1995) mention as a factor that improves reliability.

The first factor extracted from the test data was overwhelmingly dominant compared to the other factors. According to Hambleton, Swaminathan, and Rogers (1991) a dominant factor provides one with evidence to claim that a measure is unidimensional. Thus, BUSPE 2001 was considered to have met the unidimensionality assumption. This can also be interpreted as BUSPE 2001's items although classified as grammar, reading and vocabulary items measure the same construct, that is to say a composite construct.

The means of interitem correlations for groups in restricted range ability groups which were close to zero revealed that BUSPE 2001's items were locally independent. The claim of Hambleton, Swaminathan, and Rogers (1991) that unidimensionality entails local independence was justified.

The distribution of the item discrimination indices of BUSPE 2001 illustrated that they were not homogeneous. Therefore the assumption of equal discrimination indices of the one parameter IRT model could not be met.

The performance of low ability students was low on most difficult items. However, their performance on some items was not low enough to conclude that guessing had no role at all in students' performance on these items and the whole test. Nevertheless, it would be wise to consider that too strong distracters or poorly constructed items could have been influential and therefore could also explain such performance. Consequently, whether the test data met the minimum guessing assumption of the one and two parameter IRT models remained vague.

The students had been given plenty of time to complete BUSPE 2001. This was reflected in the variance of incorrect responses and omitted items the ratio of which was close to zero. It was concluded that the test data met the assumption of nonspeededness.

The chi square goodness of fit statistics suggested that the test data fit neither the one nor the two parameter IRT model when all cases were included, because the number of items that were not fitting was high and the chi square statistics were significant. When students below total scores of 30 were excluded, the number of items not fitting the model decreased in the two parameter IRT model more than it did in the one parameter model. In addition, the chi square statistic became insignificant for the two parameter IRT model. BUSPE 2001 was a multiple choice test providing four choices for each item and incorrect responses were not subject to any score reduction therefore scores below thirty were not very meaningful, or guessing could have been minimized by excluding these cases. Hence, this improvement in the goodness of fit statistics could have occurred.

The ability estimates of both the one and two parameter IRT model seemed to predict the actual scores that is the total scores of BUSPE 2001 very well, which is something expected from viable IRT models.

The invariance analyses revealed that the item parameter estimates were more invariant in the two parameter IRT model across high and low ability groups than they were in the one parameter IRT model. Yet, the item parameter estimates of the one parameter model also appeared to be highly invariant across these different groups of students. Similarly, the ability estimates of the two parameter IRT model were more invariant across odd-even and easy-difficult sets of items than those of the one parameter IRT model where the one parameter IRT model's ability estimates were also quite invariant.

Considering all goodness of fit analyses it was concluded that the two parameter IRT model was more appropriate than the one parameter IRT model for BUSPE 2001.

The predictive validity analyses showed that BUSPE 2001 had a very high classical predictive validity of .680 with DEC 1 and .572 with DEC 2 scores. These results seemed to be very high when the predictive validity criterion suggested by Alderson, Clapham, and Wall (1995) is taken into consideration.

All IRT estimated predictive validity correlations appeared to be superior to classical (total score) predictive validity correlations. In addition, all predictive validity correlations with DEC 1 scores tended to be higher than with those of the DEC 2 scores as it was in Pack's study (1968, cited in Hale et al., 1984), the proficiency exam was more related to the first English course taken than it was to the subsequent English course.

The 2P ability estimates had higher correlations with DEC 1-2 scores compared to the total scores of BUSPE 2001. The 2P ability estimates obtained by the fifty items providing highest information had slightly higher correlations with DEC 1-2 scores than the total scores of BUSPE 2001 did although only half of the items were used. Even when the thirty items providing highest information were selected, the 2P ability estimates obtained through them yielded high correlations with DEC 1-2 scores close to those of the total scores of BUSPE 2001 with DEC 1-2 scores. When items were reduced to fifty and thirty, the items were not excluded with the same proportion in each subtest. Thus, there might have been a threat towards content validity. On the other hand, construct validity was not affected, because BUSPE 2001 was found to be unidimensional, that is, all items seemed to measure more or less the same dominant construct. Moreover, reducing the number of items did not have an effect on reliability. Both sets of fifty and thirty items had high reliability coefficients, which were close to that of the whole test.

The total scores of the reading subtest was the best predictor of DEC 1-2 scores followed by the total scores of the grammar and vocabulary subtests respectively. The 2P ability estimates obtained through the subtests had higher predictive power than their total scores.

5.2 Conclusion

The conclusions of this study can briefly be summarized as follows:

1. BUSPE 2001 was highly reliable
2. BUSPE 2001 was unidimensional
3. The items of BUSPE 2001 were locally independent
4. The discrimination indices of BUSPE 2001's items were homogeneous
5. Minimal guessing was viable to some extent
6. BUSPE 2001 was nonspeeded
7. The observed distribution of BUSPE 2001's scores fit the theoretical distribution of the two parameter IRT model
8. There is a strong relationship between ability estimates of the two parameter IRT model and the actual total scores of BUSPE 2001
9. The item parameters estimated by two parameter IRT model was more invariant across high and low scoring groups

10. The ability parameters estimated by the two parameter IRT model was more invariant across sets of easy-difficult and odd-even items of BUSPE 2001
11. IRT estimated predictive validity of BUSPE 2001 and its subtests outweighed its classical predictive validity.
12. BUSPE 2001 predicted DEC 1 scores better than it did DEC 2 scores.
13. The number of items in BUSPE 2001 could be reduced to fifty without altering predictive validity and to thirty with a very slight fall in predictive validity by choosing the items providing highest information.
14. Reducing the items to fifty and thirty by choosing the items with providing high information did not change reliability.
15. The reading subtest was the best predictor of DEC 1-2 scores.

5.3 Limitations of the Study

- The DEC 1-2 grades were complicated and different for each department at BU. Each DEC grade was peculiar to its department, therefore each department's DEC grade could have been predicted in a different way by BUSPE 2001, which was not taken into consideration by this study.

- BUSPE 2001 involved items where guessing could have been influential, yet the two parameter IRT model does not consider pseudo chance factor.
- The complete data of BUSPE 2001 did not yield insignificant goodness of fit chi square statistics, because of that 58 cases were omitted and the whole test data was not used for scaling.
- As it is in most predictive validity studies of English proficiency exams all students taking the proficiency exam did not have a subsequent English grade against which their proficiency exam scores could be correlated.

5.4 Implications of the Study

Even though most findings of this study bore the fact that BUSPE 2001 was a quite successful test, further implications to improve its quality and to exploit test data were drawn out.

This study indicated that the two parameter IRT model was appropriate for scaling BUSPE 2001. Other proficiency exams either administered at BU or in other institutions which are similar in nature could be scaled by using the two parameter IRT model. However, the three parameter model and other contemporary multidimensional IRT models could also be tried out for scaling such tests.

As long as predictive validity is of primary importance for language tests, ability estimations of the two parameter IRT model could be used as performance criterion instead of raw scores, because they were observed to have predicted future performance in ESP courses more accurately and more precisely.

This study implies precious ideas for item banking studies. That is to say, item information indices and ICCs provided invaluable information about the efficiency of test items. Items providing high information in a fitted IRT model and items having ICCs with steep slopes for the ability of the test takers could be selected for item banks, because thanks to item information indices it was possible to reduce the number of items in BUSPE 2001 to fifty and even thirty. These shortened forms of BUSPE 2001 were still as effective as the whole test in predicting future performance in ESP courses especially when ability estimates in the two parameter IRT model for these short item sets were used.

Since reading items outweighed grammar and vocabulary items in predicting future performance, it would be wise to increase the weightings of reading subtests in proficiency exams.

The results of this study imply that, if students' future performance, which were DEC 1-2 grades in this study, could also be represented by IRT ability estimations, the results of the predictive validity analyses could even be higher. This issue could be another area requiring further research

REFERENCES

Aiken, L. R. (1997). Psychological testing and assessment . Boston: Allyn and Bacon.

Alderson, J., Clapham, C., & Wall, D. (1995). Language test construction and evaluation. Cambridge: Cambridge University Press.

Anastasi, A., & Urbina, S. (1997). Psychological testing. Upper Saddle River, NJ: Prentice Hall

Croanbach, L. (1990). Essentials of psychological testing. New York: Harper and Row.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Florida: Holt Rinehart and Winston, inc.

De Noble, A. F., Jung, D., & Ehrlich, S. B., (1999). "Entrepreneurıal self -efficacy: The development of a measure and its relationship to entrepreneurıal action." Retrieved December 24, 2002, from http://www.babson.edu/entrep/fer/papers99/I/I_C/IC%20Text.htm

Dooey, P. (1999). "An investigation of the predictive validity of the IELTS test as an indicator of future academic success." Retrieved December 17, 2002 from http://cea.curtin.edu.au/tlf/tlf1999/dooey.html

Enginarlar, H. (2002). Principles of language testing. Ankara: METU

Fan, X. (1998). "Item response theory and classical test theory: An empirical comparison of their item/person statistics." Educational and Psychological Measurement. 58, pp. 357-381.

Green, S. B., Salkino, N. J., & Akey, T. M. (1997). Using SPSS for Windows. (2nd ed.). New Jersey: Prentice Hall, Inc.

Hale, G., Stansfield, C.W., & Duran, R. P. (1984). "Summaries of studies involving the Test of English as a Foreign Language, 1963-1982. Research Reports" (Report 16). Princeton, NJ: Educational Testing Service.

Hambleton, R. K. & Oakland, T. (1995). International perspectives on academic assessment. Boston: Kluwer Academic Publishers.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. Newbury Park, CA: Sage publication.

Heaton, J. B. (1990). Writing English language tests. New York: Longman.

Henning, G. (1987). A guide to language testing: development, evaluation and research. Cambridge: Newbury House Publishers.

Hughes, A. (1989). Testing for language teachers. Cambridge: Cambridge University Press.

Karataş, A. G. (2001). "The use of item response theory models to scale an English proficiency test." Master Thesis. The Middle East Technical University, Department of Educational Sciences, Ankara.

Kılıç, İ. (1999). "The fit of one- two- and three- parameter models of item response theory to the student selection test of the student selection and placement center." Master Thesis. The Middle East Technical University, Department of Educational Sciences, Ankara.

Linn, R. L., & Gronlund, N. E. (2000). Measurement and assessment in teaching. Upper Saddle River, NJ: Merrill

Mislevy, R. J., & Bock, D. R. (1986). PC-BILOG: Item analysis and test scoring with binary logistic models. Scientific Software Inc.

Özkurt, S. (2002). "The fit of one-, two-, three- parameter models of item response theory to an English proficiency achievement test data." Master Thesis. The Middle East Technical University, Department of Educational Sciences, Ankara.

Patitas, W. (1989). "The quality of Khon Kaen University entrance examination tests ." Retrieved January 10, 2003 from websis.kku.ac.th/abstract/thesis/medu/mev/2532/mev320001e.html

Prapphal, K. (1990). "The relevance of language testing research in the planning of language programmes." Retrieved December 22, 2002 from http://pioneer.netserv.chula.ac.th/~pkanchan/html/testres.htm

Simner, M. L. (1999). "Postscript to the Canadian Psychological Association's position statement on the TOEFL." Retrieved December 24, 2002, from http://www.cpa.ca/documents/TOEFL.html

Stage, C. (1998). "A comparison between item analysis based on item response theory and classical test theory: a study of the SweSAT Subtest READ"**.** (Educational Measurement No 30.). Umea, Sweden: University of Umea, Department of Educational Measurement.

Weiss, D. J. & Yoes, M. E. (1991) Item response theory. In Hambleton, R. K. & Zall, J. W. (Eds). Advantages of educational and psychological testing. Boston.

APPENDICES

APPENDIX A

ITEM INFORMATION CURVES

ITEM: 0012

```
 PROBA-                                                                INFOR-
 BILITY                                                                MATION
      ------------------------------------------------------------------
  1.00|                                                      **********| 2.0000
      |                                                 *****          |
   .95|                                              ***               | 1.9000
      |                                    +        *                  |
   .90|                               +           **                   | 1.8000
      |                                          *                     |
   .85|                                                                | 1.7000
      |                                         *                      |
   .80|                                        *                       | 1.6000
      |                                     + *                        |
   .75|                                                                | 1.5000
      |                              +       *                         |
   .70|                                                                | 1.4000
      |                                     *                          |
   .65|                                                                | 1.3000
      |                                    *                           |
   .60|                                      +                         | 1.2000
      |                             +     *                            |
   .55|                                                                | 1.1000
      |                                  *                             |
   .50|                                                                | 1.0000
      |                                 *     +                        |
   .45|                                                                |  .9000
      |                            +                                   |
   .40|                                *                               |  .8000
      |                                                                |
   .35|                               *        +                       |  .7000
      |                           +                                    |
   .30|                              *                                 |  .6000
      |                             *                                  |
   .25|                                         +                      |  .5000
      |                          + *                                   |
   .20|                           *                                    |  .4000
      |                         +                +                     |  .3000
   .15|                          *                                     |  .3000
      |                        +*                 +                    |
   .10|                       **                   +                   |  .2000
      |                      *+                                        |
   .05|                   **++                      ++                 |  .1000
      |              ****+++                          +++              |
   .00|++++++++++++++++++                                ++++++++++++++|  .0000
      -+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
     -4.00   -3.00   -2.00   -1.00     .00    1.00    2.00    3.00    4.00

POINT OF MAXIMUM INFORMATION & STANDARD ERROR:       .2023    (     .0190)
VALUE OF MAXIMUM INFORMATION & STANDARD ERROR:      2.2462    (     .1563)
```

ITEM: 0094

```
 PROBA-                                                                 INFOR-
 BILITY                                                                 MATION
      ------------------------------------------------------------------
  1.00|                                                                | 2.0000
      |                                                                |
   .95|                                                                | 1.9000
      |                                                                |
   .90|                                                                | 1.8000
      |                                                                |
   .85|                                                                | 1.7000
      |                                                                |
   .80|                                                                | 1.6000
      |                                                                |
   .75|                                                                | 1.5000
      |                                                                |
   .70|                                                                | 1.4000
      |                                                                |
   .65|                                                                | 1.3000
      |                                                                |
   .60|                                                                | 1.2000
      |                                                             ***|
   .55|                                                       ******   | 1.1000
      |                                                *******         |
   .50|                                         *******                | 1.0000
      |                                  *******                       |
   .45|                            ******                              |  .9000
      |                     *******                                    |
   .40|              *******                                           |  .8000
      |      ********                                                  |
   .35|******                                                          |  .7000
      |                                                                |
   .30|                                                                |  .6000
      |                                                                |
   .25|                                                                |  .5000
      |                                                                |
   .20|                                                                |  .4000
      |                                                                |
   .15|                                                                |  .3000
      |                                                                |
   .10|                                                                |  .2000
      |                                                                |
   .05|                                                                |  .1000
      |                                                                |
   .00|++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++|  .0000
      -+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
     -4.00   -3.00   -2.00   -1.00     .00    1.00    2.00    3.00    4.00

POINT OF MAXIMUM INFORMATION & STANDARD ERROR:      1.4987    (       .2950)
VALUE OF MAXIMUM INFORMATION & STANDARD ERROR:       .0102    (       .0033)
```

APPENDIX B

RELIABILITY ANALYSIS - SCALE (ALPHA)

```
                         Mean      Std Dev     Cases

 1.  VAR00001           ,5420       ,4986      727,0
 2.  VAR00002           ,7497       ,4335      727,0
 3.  VAR00003           ,6341       ,4820      727,0
 4.  VAR00004           ,7662       ,4236      727,0
 5.  VAR00005           ,2682       ,4433      727,0
 6.  VAR00006           ,6726       ,4696      727,0
 7.  VAR00007           ,4195       ,4938      727,0
 8.  VAR00008           ,3948       ,4891      727,0
 9.  VAR00009           ,3796       ,4856      727,0
10.  VAR00010           ,3398       ,4740      727,0
11.  VAR00011           ,5007       ,5003      727,0
12.  VAR00012           ,3824       ,4863      727,0
13.  VAR00013           ,6080       ,4885      727,0
14.  VAR00014           ,7620       ,4261      727,0
15.  VAR00015           ,6314       ,4828      727,0
16.  VAR00016           ,6245       ,4846      727,0
17.  VAR00017           ,5722       ,4951      727,0
18.  VAR00018           ,3205       ,4670      727,0
19.  VAR00019           ,2889       ,4535      727,0
20.  VAR00020           ,5475       ,4981      727,0
21.  VAR00021           ,2160       ,4118      727,0
22.  VAR00022           ,3466       ,4762      727,0
23.  VAR00023           ,2957       ,4567      727,0
24.  VAR00024           ,6836       ,4654      727,0
25.  VAR00025           ,5970       ,4908      727,0
26.  VAR00026           ,8157       ,3880      727,0
27.  VAR00027           ,7029       ,4573      727,0
28.  VAR00028           ,6272       ,4839      727,0
29.  VAR00029           ,7235       ,4476      727,0
30.  VAR00030           ,5158       ,5001      727,0
31.  VAR00031           ,6121       ,4876      727,0
32.  VAR00032           ,3453       ,4758      727,0
33.  VAR00033           ,1967       ,3978      727,0
34.  VAR00034           ,6864       ,4643      727,0
35.  VAR00035           ,6094       ,4882      727,0
```

(Table Continued)

36.	VAR00036	,3604	,4804	727,0
37.	VAR00037	,3796	,4856	727,0
38.	VAR00038	,3453	,4758	727,0
39.	VAR00039	,3205	,4670	727,0
40.	VAR00040	,4869	,5002	727,0
41.	VAR00041	,5309	,4994	727,0
42.	VAR00042	,5351	,4991	727,0
43.	VAR00043	,9574	,2022	727,0
44.	VAR00044	,6066	,4888	727,0
45.	VAR00045	,6768	,4680	727,0
46.	VAR00046	,4979	,5003	727,0
47.	VAR00047	,5997	,4903	727,0
48.	VAR00048	,7029	,4573	727,0
49.	VAR00049	,5777	,4943	727,0
50.	VAR00050	,5021	,5003	727,0
51.	VAR00051	,8294	,3764	727,0
52.	VAR00052	,3563	,4792	727,0
53.	VAR00053	,4677	,4993	727,0
54.	VAR00054	,5268	,4996	727,0
55.	VAR00055	,4993	,5003	727,0
56.	VAR00056	,6190	,4860	727,0
57.	VAR00057	,5186	,5000	727,0
58.	VAR00058	,7125	,4529	727,0
59.	VAR00059	,7730	,4192	727,0
60.	VAR00060	,4773	,4998	727,0
61.	VAR00061	,3122	,4637	727,0
62.	VAR00062	,8088	,3935	727,0
63.	VAR00063	,7318	,4433	727,0
64.	VAR00064	,7139	,4523	727,0
65.	VAR00065	,4883	,5002	727,0
66.	VAR00066	,6190	,4860	727,0
67.	VAR00067	,6616	,4735	727,0
68.	VAR00068	,8212	,3835	727,0
69.	VAR00069	,4924	,5003	727,0
70.	VAR00070	,5530	,4975	727,0
71.	VAR00071	,5873	,4927	727,0
72.	VAR00072	,7909	,4069	727,0
73.	VAR00073	,4635	,4990	727,0
74.	VAR00074	,5942	,4914	727,0
75.	VAR00075	,4814	,5000	727,0
76.	VAR00076	,4966	,5003	727,0
77.	VAR00077	,5791	,4940	727,0
78.	VAR00078	,7882	,4089	727,0
79.	VAR00079	,7510	,4327	727,0
80.	VAR00080	,5365	,4990	727,0
81.	VAR00081	,3411	,4744	727,0

(Table Continued)

82.	VAR00082	,5901	,4922	727,0
83.	VAR00083	,4580	,4986	727,0
84.	VAR00084	,2930	,4554	727,0
85.	VAR00085	,8831	,3215	727,0
86.	VAR00086	,1348	,3417	727,0
87.	VAR00087	,2077	,4059	727,0
88.	VAR00088	,4718	,4995	727,0
89.	VAR00089	,3466	,4762	727,0
90.	VAR00090	,6231	,4849	727,0
91.	VAR00091	,5543	,4974	727,0
92.	VAR00092	,5475	,4981	727,0
93.	VAR00093	,5777	,4943	727,0
94.	VAR00094	,4140	,4929	727,0
95.	VAR00095	,7235	,4476	727,0
96.	VAR00096	,4608	,4988	727,0
97.	VAR00097	,3934	,4888	727,0
98.	VAR00098	,5296	,4995	727,0
99.	VAR00099	,4622	,4989	727,0
100.	VAR00100	,4402	,4967	727,0

```
      N of Cases =       727,0

Item Means          Mean    Minimum    Maximum     Range    Max/Min
Variance
                     ,5396     ,1348        ,9574        ,8226      7,1020
,0282

Item Variances       Mean    Minimum    Maximum     Range    Max/Min
Variance
                     ,2208     ,0409       ,2503        ,2095      6,1240
,0014

Inter-item
Correlations       Mean    Minimum    Maximum     Range    Max/Min
Variance
                   ,1342    -,1136        ,4763        ,5899    -4,1917
,0080
```

```
R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   (A L P H A)

Item-total Statistics

               Scale        Scale       Corrected
               Mean        Variance     Item-      Squared       Alpha
               if Item     if Item      Total      Multiple      if Item
               Deleted      Deleted     Corr.      Corr.          Deleted

VAR00001      53,4195     311,6240      ,2669         .          ,9394
VAR00002      53,2118     310,7071      ,3714         .          ,9390
VAR00003      53,3274     310,1930      ,3620         .          ,9390
VAR00004      53,1953     310,5293      ,3927         .          ,9389
VAR00005      53,6933     309,6207      ,4329         .          ,9388
VAR00006      53,2889     310,5501      ,3505         .          ,9391
VAR00007      53,5420     311,5571      ,2737         .          ,9394
VAR00008      53,5667     308,8519      ,4351         .          ,9388
VAR00009      53,5818     308,3345      ,4691         .          ,9386
VAR00010      53,6217     310,1226      ,3729         .          ,9390
VAR00011      53,4608     308,1607      ,4645         .          ,9386
VAR00012      53,5791     305,4176      ,6423         .          ,9379
VAR00013      53,3535     308,7853      ,4396         .          ,9387
VAR00014      53,1994     310,4078      ,3984         .          ,9389
VAR00015      53,3301     309,3730      ,4101         .          ,9389
VAR00016      53,3370     310,9510      ,3151         .          ,9392
VAR00017      53,3893     306,4860      ,5677         .          ,9382
VAR00018      53,6410     311,1120      ,3182         .          ,9392
VAR00019      53,6726     316,8569     -,0304         .          ,9405
VAR00020      53,4140     313,5790      ,1556         .          ,9399
VAR00021      53,7455     313,4269      ,2041         .          ,9396
VAR00022      53,6149     314,0250      ,1374         .          ,9399
VAR00023      53,6657     314,3193      ,1262         .          ,9399
VAR00024      53,2779     308,1761      ,5006         .          ,9385
VAR00025      53,3645     312,7995      ,2034         .          ,9397
VAR00026      53,1458     312,8217      ,2622         .          ,9394
VAR00027      53,2586     312,5722      ,2344         .          ,9395
VAR00028      53,3343     311,0272      ,3111         .          ,9392
VAR00029      53,2380     309,3524      ,4458         .          ,9387
VAR00030      53,4457     309,9609      ,3612         .          ,9390
VAR00031      53,3494     312,6326      ,2146         .          ,9396
VAR00032      53,6162     308,6776      ,4586         .          ,9387
VAR00033      53,7648     314,3509      ,1462         .          ,9397
VAR00034      53,2751     310,8498      ,3363         .          ,9391
VAR00035      53,3521     309,1844      ,4164         .          ,9388
VAR00036      53,6011     308,5845      ,4595         .          ,9387
```

(Table Continued)

VAR00037	53,5818	314,2960	,1185	.	,9400
VAR00038	53,6162	312,7327	,2147	.	,9396
VAR00039	53,6410	313,9797	,1434	.	,9399
VAR00040	53,4746	311,9136	,2495	.	,9395
VAR00041	53,4305	308,5596	,4425	.	,9387
VAR00042	53,4264	306,7243	,5490	.	,9383
VAR00043	53,0041	314,8306	,2369	.	,9395
VAR00044	53,3549	307,4248	,5196	.	,9384
VAR00045	53,2847	310,1433	,3767	.	,9390
VAR00046	53,4635	310,6733	,3202	.	,9392
VAR00047	53,3618	308,3276	,4648	.	,9386
VAR00048	53,2586	308,4785	,4908	.	,9386
VAR00049	53,3838	310,9971	,3058	.	,9393
VAR00050	53,4594	308,0751	,4695	.	,9386
VAR00051	53,1320	310,6465	,4359	.	,9389
VAR00052	53,6052	309,8067	,3874	.	,9389
VAR00053	53,4938	308,4184	,4507	.	,9387
VAR00054	53,4347	308,0670	,4707	.	,9386
VAR00055	53,4622	308,7310	,4317	.	,9388
VAR00056	53,3425	308,5120	,4583	.	,9387
VAR00057	53,4429	309,7016	,3762	.	,9390
VAR00058	53,2490	311,3470	,3140	.	,9392
VAR00059	53,1884	310,2110	,4189	.	,9389
VAR00060	53,4842	310,1289	,3518	.	,9391
VAR00061	53,6492	314,1123	,1365	.	,9399
VAR00062	53,1527	311,1736	,3777	.	,9390
VAR00063	53,2297	310,3728	,3842	.	,9390
VAR00064	53,2476	311,3133	,3167	.	,9392
VAR00065	53,4732	308,9604	,4186	.	,9388
VAR00066	53,3425	306,5285	,5763	.	,9382
VAR00067	53,2999	308,3480	,4811	.	,9386
VAR00068	53,1403	311,9335	,3316	.	,9392
VAR00069	53,4691	307,2108	,5195	.	,9384
VAR00070	53,4085	307,0381	,5326	.	,9384
VAR00071	53,3741	307,6642	,5013	.	,9385
VAR00072	53,1706	309,3345	,4940	.	,9386
VAR00073	53,4979	312,4817	,2177	.	,9396
VAR00074	53,3673	310,3016	,3482	.	,9391
VAR00075	53,4801	309,0902	,4113	.	,9388
VAR00076	53,4649	306,6238	,5534	.	,9383
VAR00077	53,3824	307,0767	,5343	.	,9384
VAR00078	53,1733	311,0966	,3680	.	,9390
VAR00079	53,2105	309,0782	,4802	.	,9386
VAR00080	53,4250	307,7516	,4895	.	,9385
VAR00081	53,6204	313,4204	,1742	.	,9398
VAR00082	53,3714	311,2255	,2939	.	,9393

(Table Continued)

VAR00083	53,5034	308,9611	,4200	.	,9388
VAR00084	53,6685	309,7894	,4101	.	,9389
VAR00085	53,0784	312,2211	,3737	.	,9391
VAR00086	53,8267	313,8515	,2150	.	,9395
VAR00087	53,7538	316,0343	,0258	.	,9401
VAR00088	53,4897	309,9197	,3640	.	,9390
VAR00089	53,6149	314,0443	,1363	.	,9399
VAR00090	53,3384	307,7118	,5069	.	,9385
VAR00091	53,4072	310,5613	,3287	.	,9392
VAR00092	53,4140	308,5956	,4416	.	,9387
VAR00093	53,3838	308,9834	,4226	.	,9388
VAR00094	53,5475	315,7357	,0338	.	,9403
VAR00095	53,2380	308,8813	,4761	.	,9386
VAR00096	53,5007	309,2228	,4047	.	,9389
VAR00097	53,5681	313,4826	,1646	.	,9398
VAR00098	53,4319	310,6975	,3194	.	,9392
VAR00099	53,4993	312,7490	,2025	.	,9397
VAR00100	53,5213	313,8064	,1431	.	,9399

Reliability Coefficients 100 items

Alpha = ,9396 Standardized item alpha = ,9394

APPENDIX C

PRINCIPAL COMPONENT ANALYSES

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	16,337	16,337	16,337	16,337	16,337	16,337
2	3,599	3,599	19,936	3,599	3,599	19,936
3	2,045	2,045	21,981	2,045	2,045	21,981
4	1,623	1,623	23,605			
5	1,550	1,550	25,154			
6	1,484	1,484	26,638			
7	1,439	1,439	28,077			
8	1,407	1,407	29,484			
9	1,362	1,362	30,846			
10	1,343	1,343	32,189			
11	1,321	1,321	33,510			
12	1,290	1,290	34,800			
13	1,287	1,287	36,087			
14	1,273	1,273	37,360			
15	1,255	1,255	38,614			
16	1,234	1,234	39,848			
17	1,225	1,225	41,073			
18	1,204	1,204	42,277			
19	1,177	1,177	43,455			
20	1,166	1,166	44,621			
21	1,147	1,147	45,768			
22	1,137	1,137	46,905			
23	1,112	1,112	48,017			
24	1,101	1,101	49,118			
25	1,088	1,088	50,206			
26	1,074	1,074	51,279			
27	1,057	1,057	52,337			
28	1,048	1,048	53,385			
29	1,046	1,046	54,431			
30	1,010	1,010	55,441			

(Table Continued)

31	,992	,992	56,433
32	,986	,986	57,419
33	,958	,958	58,378
34	,949	,949	59,327
35	,935	,935	60,262
36	,925	,925	61,187
37	,914	,914	62,101
38	,910	,910	63,011
39	,890	,890	63,902
40	,882	,882	64,784
41	,872	,872	65,655
42	,858	,858	66,513
43	,850	,850	67,363
44	,843	,843	68,206
45	,829	,829	69,035
46	,825	,825	69,860
47	,810	,810	70,670
48	,803	,803	71,472
49	,790	,790	72,262
50	,782	,782	73,045
51	,769	,769	73,814
52	,764	,764	74,578
53	,754	,754	75,332
54	,737	,737	76,068
55	,726	,726	76,794
56	,709	,709	77,504
57	,704	,704	78,208
58	,691	,691	78,899
59	,685	,685	79,584
60	,677	,677	80,261
61	,663	,663	80,923
62	,661	,661	81,584
63	,656	,656	82,240
64	,643	,643	82,882
65	,638	,638	83,521
66	,619	,619	84,140
67	,614	,614	84,754
68	,603	,603	85,357
69	,597	,597	85,954
70	,589	,589	86,543
71	,577	,577	87,120
72	,572	,572	87,692
73	,556	,556	88,248
74	,551	,551	88,799
75	,538	,538	89,338
76	,522	,522	89,859

(Table Continued)

77	,519	,519	90,378
78	,514	,514	90,892
79	,501	,501	91,393
80	,493	,493	91,886
81	,490	,490	92,375
82	,479	,479	92,855
83	,471	,471	93,326
84	,466	,466	93,791
85	,458	,458	94,249
86	,454	,454	94,704
87	,441	,441	95,144
88	,437	,437	95,581
89	,426	,426	96,007
90	,411	,411	96,418
91	,406	,406	96,824
92	,398	,398	97,221
93	,385	,385	97,606
94	,378	,378	97,984
95	,363	,363	98,347
96	,353	,353	98,700
97	,347	,347	99,048
98	,339	,339	99,387
99	,314	,314	99,701
100	,299	,299	100,000

Extraction Method: Principal Component Analysis.

Component Matrix

	Component		
	1	2	3
ITEM12	,693	-,165	,223
ITEM66	,626	-,169	-4,539E-02
ITEM42	,605	-,211	-1,181E-02
ITEM76	,595	-7,215E-02	6,874E-03
ITEM17	,590	,130	5,126E-02
ITEM77	,586	-,202	-,103
ITEM70	,579	-,148	1,455E-02
ITEM69	,572	-,248	6,690E-02
ITEM44	,562	-,115	2,880E-02
ITEM80	,545	-,256	-4,762E-02
ITEM71	,542	-,125	-,140
ITEM90	,539	-4,732E-03	-1,700E-02
ITEM48	,537	-,170	-,181
ITEM67	,530	-,182	-,172

(Table Continued)

ITEM24	,524	,108	-,159
ITEM72	,524	-9,021E-03	-,243
ITEM79	,523	-,128	-9,093E-02
ITEM50	,512	-,150	,106
ITEM11	,511	-,170	7,823E-02
ITEM56	,509	-,192	-,171
ITEM95	,504	6,822E-03	-5,490E-02
ITEM47	,503	-,108	-3,215E-02
ITEM9	,502	-5,578E-02	,317
ITEM54	,496	2,248E-02	2,790E-02
ITEM32	,495	-7,429E-02	,196
ITEM53	,492	-,103	5,122E-02
ITEM29	,488	-,122	-,147
ITEM8	,479	-,123	,247
ITEM13	,470	-3,854E-02	,153
ITEM41	,469	1,392E-03	-1,069E-02
ITEM55	,469	-6,885E-02	4,531E-02
ITEM59	,469	-,194	-,107
ITEM51	,466	2,762E-02	-,242
ITEM5	,462	-7,193E-02	,394
ITEM75	,456	-,107	-,166
ITEM36	,454	,306	9,257E-02
ITEM65	,451	-8,091E-02	6,266E-02
ITEM92	,447	,193	,107
ITEM15	,445	-5,079E-02	-2,528E-02
ITEM83	,443	7,735E-02	9,497E-02
ITEM4	,427	-5,437E-02	-,108
ITEM84	,423	7,197E-02	,203
ITEM14	,421	4,887E-02	-9,971E-02
ITEM52	,419	-,115	,148
ITEM93	,417	,296	8,967E-02
ITEM45	,417	-8,003E-02	1,375E-02
ITEM63	,410	-4,473E-03	-,216
ITEM62	,409	-8,910E-02	-,143
ITEM85	,403	-4,372E-02	-,247
ITEM78	,401	-,108	-6,014E-02
ITEM96	,401	,260	,124
ITEM3	,401	-,168	4,303E-02
ITEM57	,396	4,191E-02	-1,440E-02
ITEM88	,388	5,901E-03	5,165E-02
ITEM10	,385	,125	,199
ITEM30	,382	4,650E-02	8,908E-02
ITEM74	,382	-,116	4,663E-02
ITEM60	,371	5,976E-02	,167
ITEM68	,363	-6,657E-02	-,207
ITEM34	,359	4,465E-02	-2,513E-02

(Table Continued)

ITEM6	,357	,225	-7,039E-02
ITEM46	,354	-,104	7,574E-02
ITEM58	,339	-5,232E-02	-1,045E-02
ITEM64	,335	1,512E-02	-,109
ITEM49	,334	-,119	1,803E-02
ITEM98	,327	,145	-3,938E-03
ITEM16	,314	,275	-,180
ITEM82	,294	,201	-5,844E-02
ITEM1	,281	6,224E-02	,180
ITEM26	,273	,124	-7,308E-02
ITEM27	,241	7,706E-02	-4,711E-02
ITEM73	,227	5,640E-02	-,142
ITEM99	,211	4,717E-02	1,978E-02
ITEM81	,167	,166	5,173E-02
ITEM61	,145	1,749E-02	3,176E-02
ITEM21	,168	,495	-1,676E-02
ITEM35	,402	,442	-5,079E-02
ITEM25	,187	,425	-,199
ITEM2	,362	,418	-,136
ITEM97	,132	,401	-7,525E-02
ITEM23	9,768E-02	,391	6,765E-02
ITEM28	,299	,364	-4,307E-02
ITEM18	,303	,353	,106
ITEM7	,256	,348	,158
ITEM91	,315	,339	-9,173E-02
ITEM94	-1,594E-04	,333	-8,862E-02
ITEM31	,204	,307	-,143
ITEM87	-1,805E-03	,278	3,786E-02
ITEM40	,240	,273	6,978E-02
ITEM20	,142	,270	-,105
ITEM22	,123	,237	-3,654E-02
ITEM37	,104	,207	-6,797E-02
ITEM100	,138	,166	-7,076E-02
ITEM89	,137	,143	-4,657E-02
ITEM33	,143	5,016E-02	,448
ITEM86	,213	9,114E-02	,329
ITEM43	,261	-9,694E-02	-,292
ITEM39	,142	,120	,251
ITEM38	,208	,182	,230
ITEM19	-2,437E-02	-,105	,183

Extraction Method: Prinicpal Component Analysis

3 Factors Extracted

APPENDIX D

FIT OF ONE PARAMETER MODEL TO THE TEST WITH 669 CASES

```
ITEM   INTERCEPT   SLOPE    THRESHOLD  DISPERSN   ASYMPTOTE   CHISQ    DF
          S.E.       S.E.       S.E.      S.E.       S.E.      (PROB)
 ------------------------------------------------------------------------
-
 0001 |     .177  |    .443  |   -.400  |   2.258  |    .000  |    22.6  16.0
      |     .047* |    .006* |    .107* |    .028* |    .000* | ( .1248)
      |           |          |          |          |          |
 0002 |     .833  |    .443  |  -1.882  |   2.258  |    .000  |    16.0  13.0
      |     .058* |    .006* |    .134* |    .028* |    .000* | ( .2463)
      |           |          |          |          |          |
 0003 |     .442  |    .443  |   -.999  |   2.258  |    .000  |    20.7  15.0
      |     .051* |    .006* |    .118* |    .028* |    .000* | ( .1456)
      |           |          |          |          |          |
 0004 |     .903  |    .443  |  -2.038  |   2.258  |    .000  |    18.3  13.0
      |     .061* |    .006* |    .139* |    .028* |    .000* | ( .1460)
      |           |          |          |          |          |
 0005 |    -.623  |    .443  |   1.406  |   2.258  |    .000  |    30.1  16.0
      |     .057* |    .006* |    .130* |    .028* |    .000* | ( .0174)
      |           |          |          |          |          |
 0006 |     .555  |    .443  |  -1.254  |   2.258  |    .000  |    12.9  14.0
      |     .053* |    .006* |    .121* |    .028* |    .000* | ( .5310)
      |           |          |          |          |          |
 0007 |    -.166  |    .443  |    .375  |   2.258  |    .000  |    35.6  18.0
      |     .048* |    .006* |    .109* |    .028* |    .000* | ( .0081)
      |           |          |          |          |          |
 0008 |    -.240  |    .443  |    .542  |   2.258  |    .000  |    25.7  18.0
      |     .052* |    .006* |    .119* |    .028* |    .000* | ( .1074)
      |           |          |          |          |          |
 0009 |    -.290  |    .443  |    .655  |   2.258  |    .000  |    30.6  17.0
      |     .054* |    .006* |    .123* |    .028* |    .000* | ( .0223)
      |           |          |          |          |          |
 0010 |    -.393  |    .443  |    .887  |   2.258  |    .000  |    19.4  18.0
      |     .052* |    .006* |    .118* |    .028* |    .000* | ( .3685)
      |           |          |          |          |          |
 0011 |     .092  |    .443  |   -.209  |   2.258  |    .000  |    23.3  17.0
      |     .051* |    .006* |    .116* |    .028* |    .000* | ( .1379)
      |           |          |          |          |          |
 0012 |    -.257  |    .443  |    .579  |   2.258  |    .000  |    89.5  15.0
      |     .058* |    .006* |    .132* |    .028* |    .000* | ( .0000)
      |           |          |          |          |          |
 0013 |     .373  |    .443  |   -.842  |   2.258  |    .000  |    26.5  14.0
      |     .052* |    .006* |    .119* |    .028* |    .000* | ( .0222)
      |           |          |          |          |          |
 0014 |     .862  |    .443  |  -1.946  |   2.258  |    .000  |    15.4  12.0
      |     .060* |    .006* |    .138* |    .028* |    .000* | ( .2192)
      |           |          |          |          |          |
 0015 |     .447  |    .443  |  -1.009  |   2.258  |    .000  |    14.5  15.0
      |     .052* |    .006* |    .119* |    .028* |    .000* | ( .4881)
      |           |          |          |          |          |
 0016 |     .420  |    .443  |   -.949  |   2.258  |    .000  |    18.3  16.0
      |     .050* |    .006* |    .114* |    .028* |    .000* | ( .3064)
      |           |          |          |          |          |
```

(Table Continued)

0017	.301	.443	-.680	2.258	.000	43.2	14.0
	.054*	.006*	.124*	.028*	.000*	(.0001)	
0018	-.450	.443	1.017	2.258	.000	22.6	18.0
	.051*	.006*	.117*	.028*	.000*	(.2076)	
0019	-.618	.443	1.395	2.258	.000	100.2	16.0
	.048*	.006*	.111*	.028*	.000*	(.0000)	
0020	.169	.443	-.382	2.258	.000	44.6	16.0
	.046*	.006*	.105*	.028*	.000*	(.0002)	
0021	-.810	.443	1.828	2.258	.000	28.4	15.0
	.055*	.006*	.126*	.028*	.000*	(.0192)	
0022	-.393	.443	.888	2.258	.000	51.7	18.0
	.047*	.006*	.107*	.028*	.000*	(.0000)	
0023	-.537	.443	1.213	2.258	.000	55.4	17.0
	.049*	.006*	.111*	.028*	.000*	(.0000)	
0024	.631	.443	-1.425	2.258	.000	29.1	13.0
	.057*	.006*	.130*	.028*	.000*	(.0065)	
0025	.318	.443	-.717	2.258	.000	31.3	16.0
	.047*	.006*	.108*	.028*	.000*	(.0125)	
0026	1.082	.443	-2.443	2.258	.000	9.0	10.0
	.063*	.006*	.145*	.028*	.000*	(.5322)	
0027	.655	.443	-1.480	2.258	.000	13.0	14.0
	.052*	.006*	.119*	.028*	.000*	(.5267)	
0028	.425	.443	-.959	2.258	.000	13.3	16.0
	.050*	.006*	.114*	.028*	.000*	(.6522)	
0029	.758	.443	-1.711	2.258	.000	23.4	13.0
	.058*	.006*	.133*	.028*	.000*	(.0368)	
0030	.112	.443	-.254	2.258	.000	17.3	15.0
	.049*	.006*	.112*	.028*	.000*	(.2987)	
0031	.381	.443	-.861	2.258	.000	40.1	16.0
	.048*	.006*	.109*	.028*	.000*	(.0008)	
0032	-.358	.443	.809	2.258	.000	19.9	18.0
	.053*	.006*	.121*	.028*	.000*	(.3362)	
0033	-.944	.443	2.132	2.258	.000	58.2	14.0
	.060*	.006*	.137*	.028*	.000*	(.0000)	
0034	.598	.443	-1.349	2.258	.000	16.1	14.0
	.053*	.006*	.122*	.028*	.000*	(.3072)	
0035	.386	.443	-.871	2.258	.000	14.4	15.0
	.052*	.006*	.118*	.028*	.000*	(.4938)	
0036	-.311	.443	.703	2.258	.000	26.9	18.0
	.052*	.006*	.120*	.028*	.000*	(.0798)	
0037	-.299	.443	.675	2.258	.000	61.3	18.0
	.046*	.006*	.105*	.028*	.000*	(.0000)	
0038	-.376	.443	.848	2.258	.000	40.8	18.0
	.048*	.006*	.110*	.028*	.000*	(.0017)	
0039	-.486	.443	1.098	2.258	.000	45.9	18.0
	.049*	.006*	.112*	.028*	.000*	(.0003)	

(Table Continued)

```
 0040 |     .028   |     .443   |    -.063   |    2.258   |      .000   |      26.8  17.0
      |     .046*  |     .006*  |     .106*  |     .028*  |      .000*  | ( .0606)
      |            |            |            |            |             |
 0041 |     .173   |     .443   |    -.391   |    2.258   |      .000   |      15.5  16.0
      |     .050*  |     .006*  |     .116*  |     .028*  |      .000*  | ( .4908)
      |            |            |            |            |             |
 0042 |     .182   |     .443   |    -.410   |    2.258   |      .000   |      41.4  14.0
      |     .053*  |     .006*  |     .122*  |     .028*  |      .000*  | ( .0002)
      |            |            |            |            |             |
 0043 |   2.134    |     .443   |  -4.819    |    2.258   |      .000   |       5.2   1.0
      |     .131*  |     .006*  |     .299*  |     .028*  |      .000*  | ( .0219)
      |            |            |            |            |             |
 0044 |     .395   |     .443   |    -.891   |    2.258   |      .000   |      32.4  14.0
      |     .054*  |     .006*  |     .124*  |     .028*  |      .000*  | ( .0036)
      |            |            |            |            |             |
 0045 |     .583   |     .443   |  -1.317    |    2.258   |      .000   |       4.8  13.0
      |     .053*  |     .006*  |     .122*  |     .028*  |      .000*  | ( .9788)
      |            |            |            |            |             |
 0046 |     .040   |     .443   |    -.090   |    2.258   |      .000   |      11.8  17.0
      |     .048*  |     .006*  |     .111*  |     .028*  |      .000*  | ( .8142)
      |            |            |            |            |             |
 0047 |     .373   |     .443   |    -.842   |    2.258   |      .000   |      14.7  15.0
      |     .052*  |     .006*  |     .120*  |     .028*  |      .000*  | ( .4705)
      |            |            |            |            |             |
 0048 |     .701   |     .443   |  -1.582    |    2.258   |      .000   |      38.5  12.0
      |     .058*  |     .006*  |     .132*  |     .028*  |      .000*  | ( .0001)
      |            |            |            |            |             |
 0049 |     .272   |     .443   |    -.614   |    2.258   |      .000   |      19.1  16.0
      |     .049*  |     .006*  |     .111*  |     .028*  |      .000*  | ( .2634)
      |            |            |            |            |             |
 0050 |     .068   |     .443   |    -.154   |    2.258   |      .000   |      20.3  15.0
      |     .052*  |     .006*  |     .118*  |     .028*  |      .000*  | ( .1600)
      |            |            |            |            |             |
 0051 |   1.226    |     .443   |  -2.767    |    2.258   |      .000   |      14.4   7.0
      |     .072*  |     .006*  |     .165*  |     .028*  |      .000*  | ( .0436)
      |            |            |            |            |             |
 0052 |   -.333    |     .443   |     .751   |    2.258   |      .000   |      22.9  18.0
      |     .051*  |     .006*  |     .117*  |     .028*  |      .000*  | ( .1941)
      |            |            |            |            |             |
 0053 |   -.016    |     .443   |     .037   |    2.258   |      .000   |      25.8  17.0
      |     .051*  |     .006*  |     .116*  |     .028*  |      .000*  | ( .0780)
      |            |            |            |            |             |
 0054 |     .165   |     .443   |    -.373   |    2.258   |      .000   |      19.5  16.0
      |     .051*  |     .006*  |     .117*  |     .028*  |      .000*  | ( .2456)
      |            |            |            |            |             |
 0055 |     .060   |     .443   |    -.136   |    2.258   |      .000   |      18.8  16.0
      |     .051*  |     .006*  |     .116*  |     .028*  |      .000*  | ( .2767)
      |            |            |            |            |             |
 0056 |     .447   |     .443   |  -1.009    |    2.258   |      .000   |      13.5  15.0
      |     .052*  |     .006*  |     .120*  |     .028*  |      .000*  | ( .5661)
      |            |            |            |            |             |
 0057 |     .137   |     .443   |    -.309   |    2.258   |      .000   |      26.4  16.0
      |     .049*  |     .006*  |     .112*  |     .028*  |      .000*  | ( .0489)
      |            |            |            |            |             |
 0058 |     .680   |     .443   |  -1.536    |    2.258   |      .000   |       6.6  14.0
      |     .054*  |     .006*  |     .124*  |     .028*  |      .000*  | ( .9485)
      |            |            |            |            |             |
 0059 |     .926   |     .443   |  -2.092    |    2.258   |      .000   |      21.8  11.0
      |     .062*  |     .006*  |     .142*  |     .028*  |      .000*  | ( .0263)
      |            |            |            |            |             |
 0060 |     .004   |     .443   |    -.009   |    2.258   |      .000   |      10.8  17.0
      |     .049*  |     .006*  |     .111*  |     .028*  |      .000*  | ( .8689)
      |            |            |            |            |             |
 0061 |   -.505    |     .443   |   1.140    |    2.258   |      .000   |      40.1  18.0
      |     .048*  |     .006*  |     .111*  |     .028*  |      .000*  | ( .0021)
      |            |            |            |            |             |
 0062 |   1.089    |     .443   |  -2.460    |    2.258   |      .000   |      11.8  10.0
      |     .066*  |     .006*  |     .150*  |     .028*  |      .000*  | ( .2969)
      |            |            |            |            |             |
```

(Table Continued)

```
0063 |    .784  |    .443  | -1.771  |  2.258  |   .000  |    8.0  13.0
     |    .057* |    .006* |   .131* |   .028* |   .000* | ( .8471)
     |          |          |         |         |         |
0064 |    .675  |    .443  | -1.525  |  2.258  |   .000  |    9.7  13.0
     |    .054* |    .006* |   .124* |   .028* |   .000* | ( .7198)
     |          |          |         |         |         |
0065 |    .028  |    .443  |  -.063  |  2.258  |   .000  |   26.0  16.0
     |    .050* |    .006* |   .115* |   .028* |   .000* | ( .0542)
     |          |          |         |         |         |
0066 |    .425  |    .443  |  -.960  |  2.258  |   .000  |   59.6  12.0
     |    .056* |    .006* |   .128* |   .028* |   .000* | ( .0000)
     |          |          |         |         |         |
0067 |    .565  |    .443  | -1.275  |  2.258  |   .000  |   21.5  13.0
     |    .055* |    .006* |   .126* |   .028* |   .000* | ( .0635)
     |          |          |         |         |         |
0068 |   1.140  |    .443  | -2.574  |  2.258  |   .000  |   11.5  10.0
     |    .067* |    .006* |   .153* |   .028* |   .000* | ( .3189)
     |          |          |         |         |         |
0069 |    .064  |    .443  |  -.145  |  2.258  |   .000  |   28.0  15.0
     |    .052* |    .006* |   .119* |   .028* |   .000* | ( .0215)
     |          |          |         |         |         |
0070 |    .239  |    .443  |  -.539  |  2.258  |   .000  |   21.9  13.0
     |    .053* |    .006* |   .121* |   .028* |   .000* | ( .0566)
     |          |          |         |         |         |
0071 |    .322  |    .443  |  -.727  |  2.258  |   .000  |   29.5  15.0
     |    .053* |    .006* |   .122* |   .028* |   .000* | ( .0140)
     |          |          |         |         |         |
0072 |   1.062  |    .443  | -2.397  |  2.258  |   .000  |   31.0   9.0
     |    .067* |    .006* |   .154* |   .028* |   .000* | ( .0003)
     |          |          |         |         |         |
0073 |   -.045  |    .443  |   .101  |  2.258  |   .000  |   23.3  17.0
     |    .046* |    .006* |   .105* |   .028* |   .000* | ( .1392)
     |          |          |         |         |         |
0074 |    .326  |    .443  |  -.737  |  2.258  |   .000  |   11.9  16.0
     |    .050* |    .006* |   .114* |   .028* |   .000* | ( .7520)
     |          |          |         |         |         |
0075 |    .032  |    .443  |  -.072  |  2.258  |   .000  |   23.7  17.0
     |    .049* |    .006* |   .113* |   .028* |   .000* | ( .1264)
     |          |          |         |         |         |
0076 |    .076  |    .443  |  -.172  |  2.258  |   .000  |   40.3  15.0
     |    .053* |    .006* |   .122* |   .028* |   .000* | ( .0004)
     |          |          |         |         |         |
0077 |    .318  |    .443  |  -.718  |  2.258  |   .000  |   32.4  15.0
     |    .053* |    .006* |   .123* |   .028* |   .000* | ( .0057)
     |          |          |         |         |         |
0078 |    .989  |    .443  | -2.232  |  2.258  |   .000  |   11.0  10.0
     |    .062* |    .006* |   .142* |   .028* |   .000* | ( .3580)
     |          |          |         |         |         |
0079 |    .862  |    .443  | -1.946  |  2.258  |   .000  |   18.1  11.0
     |    .061* |    .006* |   .140* |   .028* |   .000* | ( .0789)
     |          |          |         |         |         |
0080 |    .198  |    .443  |  -.447  |  2.258  |   .000  |   16.0  15.0
     |    .051* |    .006* |   .118* |   .028* |   .000* | ( .3819)
     |          |          |         |         |         |
0081 |   -.402  |    .443  |   .907  |  2.258  |   .000  |   42.2  18.0
     |    .048* |    .006* |   .109* |   .028* |   .000* | ( .0011)
     |          |          |         |         |         |
0082 |    .335  |    .443  |  -.756  |  2.258  |   .000  |   34.0  16.0
     |    .048* |    .006* |   .111* |   .028* |   .000* | ( .0056)
     |          |          |         |         |         |
0083 |   -.028  |    .443  |   .064  |  2.258  |   .000  |   23.3  16.0
     |    .050* |    .006* |   .114* |   .028* |   .000* | ( .1056)
     |          |          |         |         |         |
0084 |   -.542  |    .443  |  1.223  |  2.258  |   .000  |   21.9  17.0
     |    .055* |    .006* |   .126* |   .028* |   .000* | ( .1889)
     |          |          |         |         |         |
0085 |   1.480  |    .443  | -3.342  |  2.258  |   .000  |   18.5   6.0
     |    .083* |    .006* |   .191* |   .028* |   .000* | ( .0052)
     |          |          |         |         |         |
```

(Table Continued)

```
 0086 |  -1.202  |    .443  |   2.713  |   2.258  |     .000  |    10.2  11.0
      |     .069* |    .006* |    .157* |    .028* |     .000* | ( .5107)
      |          |          |          |          |           |
 0087 |   -.872  |    .443  |   1.970  |   2.258  |     .000  |    73.1  14.0
      |     .054* |    .006* |    .124* |    .028* |     .000* | ( .0000)
      |          |          |          |          |           |
 0088 |    .004  |    .443  |   -.009  |   2.258  |     .000  |     7.3  17.0
      |     .048* |    .006* |    .111* |    .028* |     .000* | ( .9786)
      |          |          |          |          |           |
 0089 |   -.376  |    .443  |    .848  |   2.258  |     .000  |    39.7  18.0
      |     .046* |    .006* |    .106* |    .028* |     .000* | ( .0024)
      |          |          |          |          |           |
 0090 |    .447  |    .443  |  -1.009  |   2.258  |     .000  |    32.5  15.0
      |     .054* |    .006* |    .124* |    .028* |     .000* | ( .0056)
      |          |          |          |          |           |
 0091 |    .230  |    .443  |   -.520  |   2.258  |     .000  |    23.3  16.0
      |     .048* |    .006* |    .111* |    .028* |     .000* | ( .1062)
      |          |          |          |          |           |
 0092 |    .218  |    .443  |   -.493  |   2.258  |     .000  |    15.0  15.0
      |     .051* |    .006* |    .116* |    .028* |     .000* | ( .4480)
      |          |          |          |          |           |
 0093 |    .305  |    .443  |   -.689  |   2.258  |     .000  |    17.0  14.0
      |     .051* |    .006* |    .116* |    .028* |     .000* | ( .2577)
      |          |          |          |          |           |
 0094 |   -.203  |    .443  |    .459  |   2.258  |     .000  |    86.4  18.0
      |     .044* |    .006* |    .100* |    .028* |     .000* | ( .0000)
      |          |          |          |          |           |
 0095 |    .752  |    .443  |  -1.699  |   2.258  |     .000  |    17.2  12.0
      |     .059* |    .006* |    .135* |    .028* |     .000* | ( .1433)
      |          |          |          |          |           |
 0096 |   -.028  |    .443  |    .064  |   2.258  |     .000  |    14.5  15.0
      |     .049* |    .006* |    .113* |    .028* |     .000* | ( .4876)
      |          |          |          |          |           |
 0097 |   -.228  |    .443  |    .514  |   2.258  |     .000  |    62.3  18.0
      |     .045* |    .006* |    .104* |    .028* |     .000* | ( .0000)
      |          |          |          |          |           |
 0098 |    .149  |    .443  |   -.336  |   2.258  |     .000  |    11.4  16.0
      |     .048* |    .006* |    .110* |    .028* |     .000* | ( .7840)
      |          |          |          |          |           |
 0099 |   -.045  |    .443  |    .101  |   2.258  |     .000  |    32.3  17.0
      |     .046* |    .006* |    .105* |    .028* |     .000* | ( .0138)
      |          |          |          |          |           |
 0100 |   -.125  |    .443  |    .283  |   2.258  |     .000  |    58.6  17.0
      |     .045* |    .006* |    .104* |    .028* |     .000* | ( .0000)
 ---------------------------------------------------------------------------
-

     LARGEST CHANGE =      .006                                2752.51483.0
                                                              ( .0000)
```

FIT OF TWO PARAMETER MODEL TO THE TEST WITH 669 CASES

```
SUBTEST prof669 ;  ITEM PARAMETERS AFTER CYCLE   6



 ITEM   INTERCEPT   SLOPE     THRESHOLD  DISPERSN   ASYMPTOTE   CHISQ    DF
          S.E.       S.E.       S.E.       S.E.       S.E.      (PROB)
 ---------------------------------------------------------------------------
-
 0001 |    .169  |    .360  |   -.469  |   2.779  |    .000  |    14.5  15.0
      |    .021* |    .025* |    .067* |    .194* |    .000* | ( .4905)
      |          |          |          |          |          |
 0002 |    .850  |    .551  |  -1.542  |   1.815  |    .000  |    12.0  11.0
      |    .026* |    .033* |    .101* |    .110* |    .000* | ( .3607)
      |          |          |          |          |          |
 0003 |    .472  |    .619  |   -.763  |   1.616  |    .000  |    13.1  12.0
      |    .023* |    .029* |    .051* |    .077* |    .000* | ( .3627)
      |          |          |          |          |          |
 0004 |   1.024  |    .797  |  -1.285  |   1.255  |    .000  |     3.4   9.0
      |    .028* |    .035* |    .063* |    .055* |    .000* | ( .9473)
      |          |          |          |          |          |
 0005 |   -.702  |    .823  |    .853  |   1.216  |    .000  |    17.6  14.0
      |    .027* |    .031* |    .044* |    .046* |    .000* | ( .2270)
      |          |          |          |          |          |
 0006 |    .566  |    .539  |  -1.050  |   1.855  |    .000  |    13.1  12.0
      |    .023* |    .031* |    .072* |    .108* |    .000* | ( .3615)
      |          |          |          |          |          |
 0007 |   -.153  |    .347  |    .442  |   2.882  |    .000  |    16.2  15.0
      |    .021* |    .025* |    .069* |    .211* |    .000* | ( .3680)
      |          |          |          |          |          |
 0008 |   -.251  |    .801  |    .313  |   1.249  |    .000  |     7.1  15.0
      |    .024* |    .032* |    .032* |    .049* |    .000* | ( .9548)
      |          |          |          |          |          |
 0009 |   -.324  |    .936  |    .347  |   1.069  |    .000  |    19.1  14.0
      |    .025* |    .037* |    .030* |    .043* |    .000* | ( .1612)
      |          |          |          |          |          |
 0010 |   -.393  |    .565  |    .697  |   1.771  |    .000  |     8.7  16.0
      |    .023* |    .027* |    .053* |    .086* |    .000* | ( .9250)
      |          |          |          |          |          |
 0011 |    .123  |    .746  |   -.165  |   1.341  |    .000  |    16.7  14.0
      |    .023* |    .031* |    .032* |    .055* |    .000* | ( .2730)
      |          |          |          |          |          |
 0012 |   -.357  |   1.763  |    .202  |    .567  |    .000  |     5.3   9.0
      |    .032* |    .061* |    .019* |    .020* |    .000* | ( .8101)
      |          |          |          |          |          |
 0013 |    .442  |    .782  |   -.565  |   1.279  |    .000  |    19.0  11.0
      |    .024* |    .033* |    .037* |    .054* |    .000* | ( .0603)
      |          |          |          |          |          |
 0014 |    .992  |    .824  |  -1.204  |   1.213  |    .000  |     1.5   8.0
      |    .027* |    .036* |    .058* |    .053* |    .000* | ( .9920)
      |          |          |          |          |          |
 0015 |    .497  |    .693  |   -.717  |   1.443  |    .000  |     9.5  12.0
      |    .024* |    .032* |    .046* |    .067* |    .000* | ( .6580)
      |          |          |          |          |          |
 0016 |    .403  |    .389  |  -1.034  |   2.569  |    .000  |    13.7  13.0
      |    .022* |    .027* |    .090* |    .178* |    .000* | ( .3988)
      |          |          |          |          |          |
 0017 |    .435  |   1.055  |   -.412  |    .948  |    .000  |    20.0  11.0
      |    .025* |    .039* |    .027* |    .035* |    .000* | ( .0452)
      |          |          |          |          |          |
 0018 |   -.427  |    .404  |   1.057  |   2.475  |    .000  |    22.3  17.0
      |    .022* |    .026* |    .086* |    .159* |    .000* | ( .1713)
      |          |          |          |          |          |
 0019 |   -.550  |    .169  |   3.251  |   5.911  |    .000  |    31.4  17.0
      |    .022* |    .022* |    .445* |    .772* |    .000* | ( .0181)
      |          |          |          |          |          |
 0020 |    .153  |    .216  |   -.708  |   4.637  |    .000  |    16.5  16.0
      |    .021* |    .023* |    .121* |    .492* |    .000* | ( .4178)
```

(Table Continued)

```
 0021 |   -.737   |    .254   |   2.896   |   3.929   |    .000   |     17.4  16.0
      |    .024*  |    .025*  |    .303*  |    .393*  |    .000*  | ( .3581)
      |           |           |           |           |           |
 0022 |   -.351   |    .200   |   1.759   |   5.006   |    .000   |     25.5  18.0
      |    .021*  |    .023*  |    .226*  |    .570*  |    .000*  | ( .1113)
      |           |           |           |           |           |
 0023 |   -.479   |    .181   |   2.640   |   5.516   |    .000   |     19.5  17.0
      |    .022*  |    .022*  |    .344*  |    .674*  |    .000*  | ( .3005)
      |           |           |           |           |           |
 0024 |    .798   |    .965   |   -.827   |   1.036   |    .000   |      7.1  10.0
      |    .026*  |    .037*  |    .039*  |    .039*  |    .000*  | ( .7208)
      |           |           |           |           |           |
 0025 |    .288   |    .232   |  -1.243   |   4.313   |    .000   |     24.0  16.0
      |    .021*  |    .023*  |    .153*  |    .428*  |    .000*  | ( .0890)
      |           |           |           |           |           |
 0026 |   1.066   |    .471   |  -2.264   |   2.124   |    .000   |      4.3   8.0
      |    .028*  |    .038*  |    .188*  |    .173*  |    .000*  | ( .8304)
      |           |           |           |           |           |
 0027 |    .613   |    .325   |  -1.889   |   3.081   |    .000   |      2.7  13.0
      |    .023*  |    .029*  |    .178*  |    .271*  |    .000*  | ( .9986)
      |           |           |           |           |           |
 0028 |    .405   |    .381   |  -1.065   |   2.626   |    .000   |     13.8  13.0
      |    .022*  |    .027*  |    .093*  |    .187*  |    .000*  | ( .3849)
      |           |           |           |           |           |
 0029 |    .922   |    .913   |  -1.009   |   1.095   |    .000   |     11.7  10.0
      |    .027*  |    .035*  |    .045*  |    .042*  |    .000*  | ( .3045)
      |           |           |           |           |           |
 0030 |    .121   |    .535   |   -.227   |   1.868   |    .000   |      9.7  13.0
      |    .022*  |    .029*  |    .042*  |    .101*  |    .000*  | ( .7199)
      |           |           |           |           |           |
 0031 |    .347   |    .235   |  -1.472   |   4.247   |    .000   |     15.4  15.0
      |    .021*  |    .024*  |    .176*  |    .438*  |    .000*  | ( .4203)
      |           |           |           |           |           |
 0032 |   -.383   |    .767   |    .500   |   1.303   |    .000   |      6.5  15.0
      |    .024*  |    .032*  |    .037*  |    .054*  |    .000*  | ( .9701)
      |           |           |           |           |           |
 0033 |   -.894   |    .379   |   2.356   |   2.635   |    .000   |     42.2  15.0
      |    .027*  |    .025*  |    .169*  |    .175*  |    .000*  | ( .0002)
      |           |           |           |           |           |
 0034 |    .615   |    .561   |  -1.095   |   1.781   |    .000   |     15.2  12.0
      |    .024*  |    .031*  |    .072*  |    .100*  |    .000*  | ( .2319)
      |           |           |           |           |           |
 0035 |    .401   |    .556   |   -.721   |   1.799   |    .000   |     16.0  12.0
      |    .023*  |    .031*  |    .055*  |    .100*  |    .000*  | ( .1919)
      |           |           |           |           |           |
 0036 |   -.318   |    .645   |    .492   |   1.550   |    .000   |      8.4  14.0
      |    .023*  |    .030*  |    .042*  |    .072*  |    .000*  | ( .8683)
      |           |           |           |           |           |
 0037 |   -.267   |    .201   |   1.328   |   4.975   |    .000   |     17.6  18.0
      |    .021*  |    .023*  |    .181*  |    .558*  |    .000*  | ( .4792)
      |           |           |           |           |           |
 0038 |   -.344   |    .299   |   1.153   |   3.349   |    .000   |     23.9  16.0
      |    .022*  |    .024*  |    .118*  |    .274*  |    .000*  | ( .0924)
      |           |           |           |           |           |
 0039 |   -.442   |    .261   |   1.691   |   3.829   |    .000   |     18.2  18.0
      |    .022*  |    .024*  |    .175*  |    .350*  |    .000*  | ( .4437)
      |           |           |           |           |           |
 0040 |    .027   |    .296   |   -.090   |   3.374   |    .000   |      9.4  15.0
      |    .021*  |    .025*  |    .070*  |    .281*  |    .000*  | ( .8576)
      |           |           |           |           |           |
 0041 |    .202   |    .675   |   -.299   |   1.482   |    .000   |      7.6  13.0
      |    .023*  |    .030*  |    .036*  |    .066*  |    .000*  | ( .8709)
      |           |           |           |           |           |
 0042 |    .310   |   1.186   |   -.261   |    .843   |    .000   |     10.8  11.0
      |    .026*  |    .042*  |    .023*  |    .030*  |    .000*  | ( .4578)
      |           |           |           |           |           |
 0043 |   3.101   |   1.726   |  -1.797   |    .579   |    .000   |       .8   1.0
      |    .062*  |    .059*  |    .065*  |    .020*  |    .000*  | ( .3642)
      |           |           |           |           |           |
```

(Table Continued)

```
0044 |     .542  |   1.034  |   -.525  |    .967  |    .000  |     14.2  11.0
     |     .025* |    .036* |    .029* |    .034* |    .000* | ( .2232)
0045 |     .641  |    .692  |   -.926  |   1.445  |    .000  |      8.1  11.0
     |     .024* |    .034* |    .054* |    .070* |    .000* | ( .7021)
0046 |     .047  |    .516  |   -.091  |   1.939  |    .000  |      8.1  15.0
     |     .022* |    .028* |    .042* |    .104* |    .000* | ( .9202)
     |           |          |          |          |          |
0047 |     .438  |    .769  |   -.570  |   1.300  |    .000  |      5.5  12.0
     |     .024* |    .031* |    .037* |    .052* |    .000* | ( .9388)
     |           |          |          |          |          |
0048 |     .940  |   1.084  |   -.867  |    .922  |    .000  |      5.8   8.0
     |     .027* |    .035* |    .035* |    .030* |    .000* | ( .6729)
     |           |          |          |          |          |
0049 |     .270  |    .465  |   -.581  |   2.151  |    .000  |     14.7  14.0
     |     .022* |    .028* |    .057* |    .127* |    .000* | ( .3975)
     |           |          |          |          |          |
0050 |     .113  |    .881  |   -.128  |   1.135  |    .000  |      5.2  12.0
     |     .024* |    .036* |    .027* |    .046* |    .000* | ( .9493)
     |           |          |          |          |          |
0051 |    1.533  |   1.064  |  -1.440  |    .939  |    .000  |      2.9   5.0
     |     .033* |    .039* |    .056* |    .034* |    .000* | ( .7164)
     |           |          |          |          |          |
0052 |    -.336  |    .607  |    .553  |   1.646  |    .000  |     28.5  15.0
     |     .023* |    .028* |    .045* |    .076* |    .000* | ( .0188)
     |           |          |          |          |          |
0053 |     .004  |    .759  |   -.005  |   1.318  |    .000  |     12.1  14.0
     |     .023* |    .032* |    .031* |    .056* |    .000* | ( .6016)
     |           |          |          |          |          |
0054 |     .199  |    .711  |   -.279  |   1.406  |    .000  |     10.2  13.0
     |     .023* |    .030* |    .034* |    .060* |    .000* | ( .6772)
     |           |          |          |          |          |
0055 |     .087  |    .742  |   -.118  |   1.348  |    .000  |      7.6  14.0
     |     .023* |    .032* |    .031* |    .057* |    .000* | ( .9071)
     |           |          |          |          |          |
0056 |     .518  |    .771  |   -.672  |   1.297  |    .000  |     13.1  12.0
     |     .024* |    .033* |    .041* |    .055* |    .000* | ( .3651)
     |           |          |          |          |          |
0057 |     .143  |    .515  |   -.279  |   1.943  |    .000  |     26.2  14.0
     |     .022* |    .027* |    .045* |    .103* |    .000* | ( .0242)
     |           |          |          |          |          |
0058 |     .690  |    .535  |  -1.289  |   1.868  |    .000  |      2.9  12.0
     |     .024* |    .032* |    .086* |    .112* |    .000* | ( .9955)
     |           |          |          |          |          |
0059 |    1.182  |   1.043  |  -1.133  |    .959  |    .000  |     14.5   8.0
     |     .029* |    .038* |    .046* |    .035* |    .000* | ( .0683)
     |           |          |          |          |          |
0060 |     .010  |    .501  |   -.021  |   1.995  |    .000  |      9.7  15.0
     |     .022* |    .027* |    .043* |    .107* |    .000* | ( .8362)
     |           |          |          |          |          |
0061 |    -.456  |    .240  |   1.896  |   4.158  |    .000  |     22.1  18.0
     |     .022* |    .024* |    .205* |    .408* |    .000* | ( .2285)
     |           |          |          |          |          |
0062 |    1.242  |    .833  |  -1.491  |   1.200  |    .000  |      7.4   7.0
     |     .030* |    .038* |    .072* |    .055* |    .000* | ( .3911)
     |           |          |          |          |          |
0063 |     .834  |    .648  |  -1.287  |   1.542  |    .000  |     10.3  11.0
     |     .026* |    .033* |    .073* |    .079* |    .000* | ( .5014)
     |           |          |          |          |          |
0064 |     .691  |    .555  |  -1.245  |   1.802  |    .000  |      9.1  11.0
     |     .024* |    .032* |    .081* |    .105* |    .000* | ( .6165)
     |           |          |          |          |          |
0065 |     .048  |    .703  |   -.069  |   1.423  |    .000  |     22.1  14.0
     |     .023* |    .031* |    .032* |    .063* |    .000* | ( .0755)
     |           |          |          |          |          |
0066 |     .740  |   1.444  |   -.512  |    .693  |    .000  |     23.8   9.0
     |     .028* |    .043* |    .023* |    .021* |    .000* | ( .0047)
     |           |          |          |          |          |
0067 |     .715  |    .948  |   -.754  |   1.055  |    .000  |     12.6  11.0
     |     .026* |    .035* |    .037* |    .039* |    .000* | ( .3196)
     |           |          |          |          |          |
```

(Table Continued)

```
0068 |   1.204  |    .659  | -1.828  |   1.518  |    .000  |     8.4   8.0
     |    .030* |    .036* |   .106* |    .084* |    .000* | ( .3974)
     |          |          |         |          |          |
0069 |    .122  |   1.000  |  -.122  |   1.000  |    .000  |    12.4  12.0
     |    .025* |    .039* |   .025* |    .039* |    .000* | ( .4159)
     |          |          |         |          |          |
0070 |    .351  |   1.046  |  -.336  |    .956  |    .000  |     8.8  10.0
     |    .025* |    .039* |   .026* |    .035* |    .000* | ( .5549)
     |          |          |         |          |          |
0071 |    .434  |    .970  |  -.447  |   1.031  |    .000  |    12.6  12.0
     |    .025* |    .035* |   .029* |    .037* |    .000* | ( .3991)
     |          |          |         |          |          |
0072 |   1.483  |   1.270  | -1.168  |    .787  |    .000  |     7.4   6.0
     |    .032* |    .038* |   .040* |    .023* |    .000* | ( .2868)
     |          |          |         |          |          |
0073 |   -.040  |    .278  |   .143  |   3.597  |    .000  |    19.7  17.0
     |    .021* |    .023* |   .075* |    .303* |    .000* | ( .2912)
     |          |          |         |          |          |
0074 |    .343  |    .577  |  -.595  |   1.733  |    .000  |    12.2  13.0
     |    .022* |    .030* |   .048* |    .091* |    .000* | ( .5118)
     |          |          |         |          |          |
0075 |    .044  |    .586  |  -.074  |   1.706  |    .000  |    27.9  15.0
     |    .022* |    .026* |   .038* |    .076* |    .000* | ( .0221)
     |          |          |         |          |          |
0076 |    .144  |   1.048  |  -.138  |    .954  |    .000  |     7.5  12.0
     |    .025* |    .038* |   .024* |    .035* |    .000* | ( .8233)
     |          |          |         |          |          |
0077 |    .458  |   1.070  |  -.428  |    .935  |    .000  |     7.2  11.0
     |    .026* |    .038* |   .027* |    .033* |    .000* | ( .7801)
     |          |          |         |          |          |
0078 |   1.115  |    .793  | -1.405  |   1.260  |    .000  |    10.7   8.0
     |    .028* |    .038* |   .071* |    .060* |    .000* | ( .2193)
     |          |          |         |          |          |
0079 |   1.204  |   1.210  |  -.995  |    .826  |    .000  |     4.0   8.0
     |    .029* |    .038* |   .036* |    .026* |    .000* | ( .8579)
     |          |          |         |          |          |
0080 |    .267  |    .894  |  -.298  |   1.119  |    .000  |     8.6  12.0
     |    .024* |    .037* |   .029* |    .046* |    .000* | ( .7392)
     |          |          |         |          |          |
0081 |   -.363  |    .241  |  1.503  |   4.143  |    .000  |    25.9  18.0
     |    .021* |    .023* |   .170* |    .403* |    .000* | ( .1027)
     |          |          |         |          |          |
0082 |    .315  |    .347  |  -.909  |   2.886  |    .000  |    31.2  14.0
     |    .022* |    .027* |   .092* |    .224* |    .000* | ( .0054)
     |          |          |         |          |          |
0083 |   -.018  |    .604  |   .029  |   1.656  |    .000  |    14.7  14.0
     |    .022* |    .030* |   .037* |    .083* |    .000* | ( .4012)
     |          |          |         |          |          |
0084 |   -.578  |    .708  |   .816  |   1.412  |    .000  |    15.4  15.0
     |    .025* |    .031* |   .050* |    .063* |    .000* | ( .4201)
     |          |          |         |          |          |
0085 |   2.033  |   1.355  | -1.501  |    .738  |    .000  |     7.6   4.0
     |    .040* |    .043* |   .051* |    .023* |    .000* | ( .1078)
     |          |          |         |          |          |
0086 |  -1.182  |    .474  |  2.492  |   2.109  |    .000  |    12.0  11.0
     |    .031* |    .028* |   .159* |    .126* |    .000* | ( .3608)
     |          |          |         |          |          |
0087 |   -.777  |    .153  |  5.077  |   6.532  |    .000  |    21.0  15.0
     |    .025* |    .023* |   .772* |    .971* |    .000* | ( .1360)
     |          |          |         |          |          |
0088 |    .011  |    .512  |  -.021  |   1.953  |    .000  |     7.9  15.0
     |    .022* |    .028* |   .042* |    .108* |    .000* | ( .9296)
     |          |          |         |          |          |
0089 |   -.335  |    .186  |  1.795  |   5.362  |    .000  |    11.1  18.0
     |    .021* |    .022* |   .241* |    .639* |    .000* | ( .8921)
     |          |          |         |          |          |
0090 |    .568  |    .920  |  -.617  |   1.087  |    .000  |     5.6  11.0
     |    .025* |    .035* |   .034* |    .041* |    .000* | ( .9003)
     |          |          |         |          |          |
```

(Table Continued)

```
 0091 |     .221  |     .384  |    -.575  |   2.601  |     .000  |     14.9  14.0
      |     .021* |     .028* |     .068* |    .188* |     .000* | ( .3866)
 0092 |     .240  |     .616  |    -.391  |   1.624  |     .000  |     12.4  12.0
      |     .022* |     .030* |     .040* |    .080* |     .000* | ( .4116)
      |           |           |           |          |           |
 0093 |     .321  |     .572  |    -.562  |   1.750  |     .000  |     13.2  11.0
      |     .022* |     .032* |     .048* |    .097* |     .000* | ( .2768)
      |           |           |           |          |           |
 0094 |    -.178  |     .119  |    1.499  |   8.414  |     .000  |     15.3  18.0
      |     .020* |     .019* |     .295* |   1.347* |     .000* | ( .6451)
      |           |           |           |          |           |
 0095 |     .964  |    1.016  |    -.948  |    .984  |     .000  |      3.3   8.0
      |     .027* |     .037* |     .041* |    .035* |     .000* | ( .9146)
      |           |           |           |          |           |
 0096 |    -.020  |     .536  |     .038  |   1.866  |     .000  |     14.0  14.0
      |     .022* |     .029* |     .041* |    .101* |     .000* | ( .4485)
      |           |           |           |          |           |
 0097 |    -.203  |     .184  |    1.101  |   5.435  |     .000  |     17.4  17.0
      |     .021* |     .022* |     .173* |    .653* |     .000* | ( .4295)
      |           |           |           |          |           |
 0098 |     .145  |     .398  |    -.364  |   2.514  |     .000  |     17.7  15.0
      |     .021* |     .025* |     .058* |    .160* |     .000* | ( .2783)
      |           |           |           |          |           |
 0099 |    -.040  |     .268  |     .148  |   3.726  |     .000  |     12.6  17.0
      |     .021* |     .024* |     .078* |    .327* |     .000* | ( .7632)
      |           |           |           |          |           |
 0100 |    -.112  |     .219  |     .513  |   4.570  |     .000  |     20.4  17.0
      |     .021* |     .023* |     .108* |    .475* |     .000* | ( .2553)
 ---------------------------------------------------------------------------
-

     LARGEST CHANGE =      .007                                1345.91278.0
                                                              ( .0911)
```

APPENDIX E

TEST INFORMATION CURVE OBTAINED FROM THE ONE PARAMETER MODEL WITH ALL (100) ITEMS AND 669 CASES WITH TOTALS ABOVE 30

```
 STANDARD                                                         INFORMATION
ERROR
      ------------------------------------------------------------------
   .46|  *                       ++++++                       *        |12.0588
      |   *                    ++      ++                    *         |
   .44|    *                 ++          ++                            |11.4558
      |                     +              +                *          |
   .41|     *              +                +              *           |10.8529
      |      *            +                  +            *            |
   .39|       *          +                    +          *             |10.2500
      |        **       +                      +        *              |
   .37|          *     +                        +     **               | 9.6470
      |           *   +                          +   *                 |
   .35|            **+                            +**                  | 9.0441
      |              **                           *+                   |
   .32|             +  **                       **                     | 8.4411
      |            +     **                  ***    +                  |
   .30|           +        ****          ****        +                 | 7.8382
      |          +             **********             +                |
   .28|                                                +               | 7.2353
      |         +                                                      |
   .25|        +                                        +              | 6.6323
      |       +                                          +             |
   .23|      +                                            +            | 6.0294
      |     +                                              +           |
   .21|    +                                                +          | 5.4264
      |   +                                                  +         |
   .18|  +                                                    +        | 4.8235
      | +                                                      +       |
   .16|+                                                        +      | 4.2206
      |                                                          +     |
   .14|                                                           +    | 3.6176
      |                                                            +   |
   .12|                                                             ++ | 3.0147
      |                                                               +|
   .09|                                                                | 2.4118
      |                                                                |
   .07|                                                                | 1.8088
      |                                                                |
   .05|                                                                | 1.2059
      |                                                                |
   .02|                                                                |  .6029
      |                                                                |
   .00|                                                                |  .0000
      -+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
     -4.00   -3.00   -2.00   -1.00     .00    1.00    2.00    3.00    4.00
```

MAXIMUM INFORMATION APPROXIMATELY   .1206D+02 AT   -1.9286

TEST INFORMATION CURVE OBTAINED FROM THE TWO PARAMETER MODEL WITH ALL (100) ITEMS AND 669 CASES WITH TOTALS ABOVE 30

```
TEST:   prof669

  STANDARD                                                              INFORMATION
  ERROR
       ------------------------------------------------------------------
 .3D+00|                          ++++               *                  |30.2837
       |           *             +    +                                 |
 .3D+00|                        +      +            *                   |28.7695
       |                       +                                        |
 .3D+00|            *                   +          *                    |27.2553
       |                      +          +                              |
 .3D+00|             *       +                    *                     |25.7411
       |                                         *                      |
 .3D+00|              *     +             +                             |24.2270
       |                                        *                       |
 .3D+00|               *   +               +                            |22.7128
       |                                       *                        |
 .2D+00|                *                   + *                         |21.1986
       |                 *+                                             |
 .2D+00|                                     *                          |19.6844
       |                 +*                 *+                          |
 .2D+00|                   **              *                            |18.1702
       |                +    *           **   +                         |
 .2D+00|                      ***      **                               |16.6560
       |                         ******        +                        |
 .2D+00|               +                                                |15.1419
       |                                                                |
 .2D+00|              +                         +                       |13.6277
       |                                         +                      |
 .1D+00|                                                                |12.1135
       |             +                            +                     |
 .1D+00|                                                                |10.5993
       |            +                              +                    |
 .1D+00|           +                                +                   | 9.0851
       |                                             +                  |
 .9D-01|          +                                   +                 | 7.5709
       |         +                                     +                |
 .7D-01|        +                                       +               | 6.0567
       |       +                                         +              |
 .5D-01|      +                                           ++            | 4.5426
       |    ++                                              ++          |
 .3D-01|  ++                                                  +++       | 3.0284
       |++                                                       ++++   |
 .2D-01|                                                             +++| 1.5142
       |                                                                |
 .0D+00|                                                                |  .0000
       -+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
      -4.00   -3.00   -2.00   -1.00     .00    1.00    2.00    3.00    4.00
```

MAXIMUM INFORMATION APPROXIMATELY .3028D+02 AT -2.0714

TEST INFORMATION CURVE OBTAINED FROM 2 PARAMETER MODEL FOR 50 ITEMS PROVIDING HIGHEST INFORMATION 669 CASES WITH TOTAL ABOVE 30

```
TEST:   inf50

  STANDARD                                                              INFOR-
  ERROR                                                                 MATION
       ------------------------------------------------------------------
    .43|         *                 ++++              *                  |16.9637
       |                         ++    +                                |
    .41|                        +       +           *                   |16.1155
       |          *            +         +                              |
    .39|                                  +        *                    |15.2673
       |           *          +                                         |
    .37|                     +             +      *                     |14.4191
       |            *       +                                           |
    .35|             *                      +    *                      |13.5710
       |                   +                                            |
    .32|              *                         *                       |12.7228
       |               *  +                  + *                        |
    .30|                **                                              |11.8746
       |                 *                    +                         |
    .28|                + **                 *                          |11.0264
       |                    *               *                           |
    .26|               +     **           **   +                        |10.1782
       |                       *****  ****                              |
    .24|              +             **          +                       | 9.3300
       |                                                                |
    .22|             +                           +                      | 8.4819
       |                                                                |
    .19|            +                             +                     | 7.6337
       |                                                                |
    .17|           +                               +                    | 6.7855
       |                                                                |
    .15|          +                                 +                   | 5.9373
       |         +                                   +                  |
    .13|                                                                | 5.0891
       |        +                                     +                 |
    .11|       +                                       +                | 4.2409
       |                                                +               |
    .09|      +                                          +              | 3.3927
       |     +                                            +             |
    .06|    +                                              +            | 2.5446
       |  ++                                                ++          |
    .04|++                                                    ++        | 1.6964
       |                                                        ++      |
    .02|                                                          ++++  |  .8482
       |                                                              ++|
    .00|                                                                |  .0000
      -+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
      -4.00   -3.00   -2.00   -1.00     .00    1.00    2.00    3.00    4.00
```

MAXIMUM INFORMATION APPROXIMATELY   .1696D+02 AT   -2.0000

TEST INFORMATION CURVE OBTAINED FROM 2 PARAMETER MODEL FOR 30 ITEMS PROVIDING HIGHEST INFORMATION 669 CASES WITH TOTAL ABOVE 30

```
TEST:   inf30

  STANDARD                                                              INFOR-
  ERROR                                                                 MATION
       ------------------------------------------------------------------
    .49|                          +++++           *                     |12.6939
       |                         +     +                                |
    .47|          *             +       +                               |12.0592
       |                       +         +       *                      |
    .44|           *          +                                         |11.4245
       |                     +            +                             |
    .42|            *                           *                       |10.7898
       |                    +                                           |
    .39|             *     +               +   *                        |10.1551
       |                                                                |
    .37|              *   +                 + *                         | 9.5204
       |               *                                                |
    .35|                *+                   *                          | 8.8857
       |                 *                                              |
    .32|                + **                *+                          | 8.2510
       |               +    *              *                            |
    .30|                     ***         **                             | 7.6163
       |              +         *********     +                         |
    .27|                                                                | 6.9816
       |                                                                |
    .25|             +                         +                        | 6.3469
       |                                                                |
    .22|            +                                                   | 5.7122
       |                                        +                       |
    .20|           +                                                    | 5.0775
       |                                         +                      |
    .17|          +                                                     | 4.4429
       |         +                                +                     |
    .15|                                                                | 3.8082
       |        +                                  +                    |
    .12|                                            +                   | 3.1735
       |       +                                                        |
    .10|      +                                      +                  | 2.5388
       |     +                                        +                 |
    .07|    +                                          +                | 1.9041
       |   +                                            +               |
    .05|  +                                              ++             | 1.2694
       |++                                                 ++           |
    .02|                                                     +++        |  .6347
       |                                                        +++++++ |
    .00|                                                               +|  .0000
       -+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
      -4.00   -3.00   -2.00   -1.00    .00     1.00    2.00    3.00    4.00
```

MAXIMUM INFORMATION APPROXIMATELY   .1269D+02 AT   -2.0000