CONSTRUCT VALIDITY AND FACTOR STRUCTURE OF STUDENT SELECTION EXAMINATION ACROSS SUBGROUPS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

SERKAN ARIKAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
SECONDARY SCIENCE AND MATHEMATICS EDUCATION

MAY 2010

Approval of the thesis:

CONSTRUCT VALIDITY AND FACTOR STRUCTURE OF STUDENT SELECTION EXAMINATION ACROSS SUBGROUPS

submitted by **SERKAN ARIKAN** in partial fulfillment of the requirements for the degree of the **Doctor of Philosophy in Secondary Science and Mathematics Education, Middle East Technical University** by,

Prof. Dr. Canan Özgen ____________
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Ömer Geban ____________
Head of Department, **Secondary Science and Mathematics Education**

Prof. Dr. Giray Berberoğlu ____________
Supervisor, **Secondary Science and Mathematics Education Dept., METU**

Examining Committee Members:

Prof. Dr. Feza Korkusuz ____________
Physical Education and Sports Dept., METU

Prof. Dr. Giray Berberoğlu ____________
Secondary Science and Mathematics Education Dept., METU

Prof. Dr. Ömer Geban ____________
Secondary Science and Mathematics Education Dept., METU

Assoc. Prof. Dr. Levent İnce ____________
Physical Education and Sports Dept., METU

Assoc. Prof. Dr. Nukhet Demirtaşlı ____________
Educational Sciences Dept., AU

Date: 11 May 2010

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.

Name, Last name: Serkan Arıkan

Signature :

ABSTRACT

CONSTRUCT VALIDITY AND FACTOR STRUCTURE OF STUDENT SELECTION EXAMINATION ACROSS SUBGROUPS

Arıkan, Serkan

Ph.D., Department of Secondary Science and Mathematics Education

Supervisor: Prof. Dr. Giray Berberoğlu

May 2010, 146 pages

In developing countries, there is a great demand for university education. In order to select students to universities a standardized test score is used. In Turkey, the Student Selection Test (SST) have important role in admission to universities. However, there is very limited knowledge about what SST mathematics sections actually measures.

The main purpose of the present study is to evaluate the content of the mathematics subtest of the SST in line with mathematical cognitive skills and eventually provide construct related evidence for dimensionality of the test items. Within this framework, it is aimed to cross validate the mathematics subtest across gender groups, school types and two consecutive years. Also relations among mathematical abilities are investigated. This study is first in investigating what is measured by SST Mathematics sections and analyzing construct validity by testing several nested confirmatory factor models.

Comparison of fit indices of five competitive models showed three-factor model has better fit indices in which Basic Computation Ability, Advanced Computation Ability and Geometry Ability is measured. It is concluded that problem solving items are not measuring a different process, but measures some sort of computation ability. There is a problem related to the content of the mathematics subtests of the SST in line with mathematical cognitive skills. Higher order cognitive skills are not measured properly.

Three-factor model is tested about the invariance of the factors across gender, school types and years. It is concluded that invariant factor structure indicates that SST mathematics section is operating similarly for subgroups and years.

The relations among mathematical abilities on three-factor model are investigated by item mapping and structural equation models. It is seen that Basic Computation Ability is a prerequisite to acquire Geometry Ability and Advanced Computation Ability.

Keywords: Construct Validity, Confirmatory Factor Analysis, Model Testing, Invariance of Factor Structures, Structural Equation Modeling, Student Selection Test (SST).

ÖZ

ÖĞRENCİ SEÇME SINAVININ YAPI GEÇERLİĞİ VE GRUPLAR ARASI FAKTÖR YAPILARININ İNCELENMESİ

Arıkan, Serkan

Doktora., Ortaöğretim Fen ve Matematik Alanları Eğitimi Bölümü

Tez Yöneticisi: Prof. Dr. Giray Berberoğlu

Mart 2010, 146 sayfa

Gelişmekte olan ülkelerde üniversite eğitimine büyük bir talep vardır. Öğrencileri üniversitelere seçmek için standart test puanları kullanılmaktadır. Türkiye'de Öğrenci Seçme Sınavı (ÖSS) sonuçları üniversiteye girişte önemli bir yer tutmaktadır. Fakat, ÖSS matematik bölümlerinin ne ölçtüğü hakkında detaylı bilgiler bulunmamaktadır.

Bu çalışmanın ana amacı ÖSS matematik bölümlerinin içeriğinin düşünme süreçlerine bağlı olarak değerlendirilmesi ve ölçülen yapıların ortaya çıkarılmasıdır. Bu çerçevede; ortaya çıkarılan yapıların cinsiyet, okul türleri ve yıllar bakımından benzerliklerinin incelenmesi amaçlanmaktadır. Ayrıca matematik becerileri arasındaki ilişkiler de incelenmektedir. Bu çalışma ÖSS matematik bölümlerinde ne ölçüldüğünü ve yapı geçerliğini birbirleri ile ilişkili birçok modeli doğrulayıcı faktör analizi ile test eden ilk çalışmadır.

Test edilen beş modelden elde edilen uyum indislerinin karşılaştırılması sonucunda, üç faktörden oluşan modelin en iyi model olduğu kabul edilmiştir. Bu modele göre Temel İşlem Becerisi, İleri İşlem Becerisi ve Geometri Becerisi ölçülmektedir. Problem çözme sorularının farklı bir düşünme süreci ölçemediği ve işlem becerisinden farklılık gösteremediği ortaya konmuştur. ÖSS matematik bölümlerinin içeriğinin düşünme süreçlerine göre problemli olduğu ve üst düzey düşünme süreçlerinin ölçülmediği belirlenmiştir.

Üç faktörlü modelin cinsiyet, okul türleri ve yıllar bakımından eşitliği incelenmiş ve faktör yapıları bakımından ÖSS'nin bu gruplar ve yıllar açısından benzerlik gösterdiği görülmüştür.

Matematik becerileri arasındaki ilişki madde haritalama yöntemi ve yapısal eşitlik modeli ile test edildiğinde Temel İşlem Becerisinin, İleri İşlem Becerisi ve Geometri Becerisinin edinilmesi için ön koşul olduğu görülmüştür.

Anahtar Kelimeler: Yapı Geçerliği, Doğrulayıcı Faktör Analizi, Modellerin Test Edilmesi, Faktör Yapılarının Eşitliği, Yapısal Eşitlik Modeli, Öğrenci Seçme Sınavı (ÖSS).

To İdil and Karya

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Dr. Giray Berberoğlu for his guidance and contributions and for being a great academic model for me.

Also, I would like to thank Prof. Dr. Ünal Yarımağan, President of ÖSYM, for giving permission to use such a great data.

Besides that, I would like to thank TUBİTAK for giving scholarship during my PHD.

Finally, I would like to thank Ömer Ahmet Konak, General Manager of Cito Turkey, for his great understanding and support during my thesis.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In developing countries, there is a great demand for university education. The higher education is the most important stage of educational path of an individual, since a university degree brings higher probability of getting a job (Tansel, 1999). Tansel (1999) further stated that as the level of education rises for an individual, private economical return to that person also increases. Therefore, higher education means welfare for many individuals. Besides that, university education not only provides job opportunities but also social status and prestige in the community. However, there isn't enough quota in universities due to size of the young population in the country. Almost always there are more applicants for a university program than available quotas. According to the statistics from The Student Selection and Placement Center of Turkey, in 2006, 1 537 377 students applied for the university admission and only 176 194 of them were admitted to four or more year programs of universities. Therefore, only 11.46% of the students had opportunities to be admitted to these programs (Student Selection and Placement Center, 2006a). In 2009, 1 451 350 students applied for the exam and only 290 097 of them were admitted to four or more year programs of universities. As it is seen from the examples, almost 20% selection ratio is too low and makes the examination very prominent in public (Student Selection and Placement Center, 2009a).

This great demand to university education makes selection indispensible. Actually, almost all countries in the world somehow select students for higher education. Although each country has its own strategies and admission rules, there is a common criterion for most countries: a standardized test score. In the United States of America, criteria used are scores on standardized tests such as The Scholastic Aptitude test (SAT) or American College Testing Program (ACT), personal statements, school references, and some other documents required by colleges. In Japan, students who want to enter a university must pass a college entrance exam. Because the results of exams is related to their future, and better universities means better job opportunities, students often think of their whole schooling as training for entrance examinations (Vernille, 2001). In France, in order to go to a university, a student must have "baccalaureat", which is the secondary school diploma that can be obtained by passing a very difficult national examination at the end of 12th grade (Vernille, 2001). In England standard attainment tasks and tests (SATs) are used to check whether students have reached the National Curriculum learning targets before age 16. At the age of 16, students in England take examinations of English, mathematics, science, and a range of elective subjects in order to earn the General Certificate of Secondary Education (GCSE). If a student passes these examinations, he or she may pursue further education by taking other examinations, such as A-level or The National Council for Vocational Qualification (Gregory and Clarke, 2003). In Turkey, admission to higher education is based on a centrally administered examination system. Since 1974, the admission of students to higher education has been carried out on the basis of results of examinations organized by The Student Selection and Placement Center (SSPC) which was affiliated with The Higher Education Council (YÖK). The basis of this system is The Student Selection Test (SST). It is stated that the aim of SST is "to select and place students with the highest probability of success in all the available higher education programs, taking into consideration their preferences, and performance on SST" (Student Selection and Placement Center, 2006b).

As it is summarized above, standardized tests have crucial role in selection decisions of universities. When SST in Turkey and other standardized tests described above are compared, there are three major differences: all items in SST are only multiple choice items; new forms of SST are prepared for each year; and selection ratio in Turkey is very low.

In all of the standardized test programs, mathematics constitutes an important part when the content of the tests are considered. For instance, in SST, two sections of mathematics tests are used; in SAT, mathematics section is one of the three major sections; in ACT, mathematics section is one of the four major sections; and in GCSE, mathematics section is one of the three major sections. Mathematics assessment is important subject which is closely related to areas such as engineering, finance, statistics, natural science or medicine. Therefore, success in mathematics is a prerequisite to be successful in many areas. Besides that, skills acquired in mathematics will probably help people to be successful in their daily life challenges. These challenges can be situations people face when they shopping, travelling, cooking, dealing with their personal finances, or judging political issues. In all of these situations, quantitative abilities, spatial reasoning ability or other mathematical competencies is necessary to clarify, formulate or solve problems (OECD, 2003).

Many standardized tests define what they aim to measure in terms of mathematical cognitive processes. For example, in TIMSS, cognitive skills measured in mathematics sections are classified as three main domains: *knowing*, which covers the facts, procedures, and concepts students need to know; *applying*, which focuses on the ability of students to apply knowledge and conceptual understanding to solve problems or answer questions; *reasoning*, which goes beyond the solution of routine problems to encompass unfamiliar situations, complex contexts, and multi-step problems (Mullis et al. 2007). In PISA, cognitive skills measured in mathematics sections are classified as three main clusters: *reproduction cluster, connections* *cluster*, and *reflection cluster*. *Reproduction cluster* includes standard representations and definitions, routine computations, routine procedures, routine solving problem. *Connections cluster* includes modeling, standard problem solving translation and interpretation, multiple well-defined methods. *Reflection cluster* includes complex problem solving and posing, reflection and insight, original mathematical approach, multiple complex methods and generalization (OECD, 2003). Similarly Bloom et al. (1971) classify cognitive processes for mathematics as four main domains: *Computation*, which includes knowledge of specific facts, knowledge of terminology, ability to carry out algorithms; *Comprehension*, which includes knowledge of concepts, knowledge of principles, rules and generalizations, knowledge of mathematical structure, ability to transform problem elements from one mode to another, ability to follow a line of reasoning, ability to read and interpret a problem; *Application*, which includes ability to solve routine problems, ability to make comparisons, ability to analyze data, ability to recognize patterns isomorphisms, and symmetries; and *Analysis*, which includes ability to solve non-routine problems, ability to discover relationships, ability to construct proofs, ability to criticize proofs, ability to formulate and validate generalizations. All of these cognitive process classifications in mathematics have common properties in which there is a hierarchical structure among these cognitive skills. These cognitive processes constitute assessment framework for many standardized tests in the world, such as PISA and TIMSS.

Skills to be assessed in large scale testing program in line with mathematics constitute the major issue of content validity. The content specifications as explained above are related to the validity of the test content. Content validity is related determining whether content of an instrument is an adequate sample of the domain of the content (Fraenkel & Wallen, 2005). The content validity per se is not enough for a defensible test score. When it is asked what an instrument really measures, information on construct validity is a necessity (Cronbach, 1971). Therefore, it can be proven that measurement instrument is indeed reflecting the construct that is considered to underlie the measure. Construct validation is defined as "research process by which one goes about establishing construct validity; that is, the process of collecting evidence that a test or other operational measure does indeed reflect the theoretical construct" (Arvey, 1992).

Construct related evidence is rather a set of statistical analyses to collect evidence whether test measures what it intended to measure. The Structural equation modeling (SEM) methods, like confirmatory factor analyses (CFA), are central for the construct validation research (Zumbo, 2005). Confirmatory factor analysis and structural equation modeling analysis are used to examine what is measured by a test, what are the dimensions of a test and what are the relations among skills measured in these tests. In many construct related validation attempts justifying the content related framework through exploratory or confirmatory analyses are general methods being used in the validation studies (Zumbo, 2005). Beside these studies, it is also important to cross validate the test scores across groups of interests, such as gender. The point in cross validation is to evaluate validation results with new sample. By doing so, validity and generalizability of results will be supported (Treat and Weersing, 2005).

As it is explained, in Turkey, the Student Selection Test (SST) is used for selection process. However, there is very limited knowledge about what SST mathematics sections actually measure. Student Selection and Placement Center only stated that first section of mathematics has items about "power of using mathematical relations" and second section of mathematics has items about "Mathematics and Geometry" (Student Selection and Placement Center, 2006e). Besides that, there is limited research about construct validity of SST. The aim of the first section implies cognitive skills which require reasoning rather than mere memorized and algorithmic calculation. The second section rather focuses on curriculum based learning outcomes.

The limited studies about the validity of SST examination make content-wise evaluation and construct validity analysis of SST mathematics sections worth investigating. This analysis should base on the theoretical framework that the SST mathematics subtests underline and the empirical evidence supporting that framework in assessing mathematic achievement of students. Thus, any attempt to validate the mathematics subtests of the SST should start with the assessment of theoretical framework underlying the content specifications and empirically justifying the dimensionality of these specifications.

1.1 Purpose of the Study

The purpose of the present study is to evaluate the content of the mathematics subtest of the SST in line with mathematical cognitive skills and eventually provide construct related evidence for dimensionality of the test items. In terms of content specifications, items in the mathematics subtest will be evaluated with reference to the mathematical cognitive processes. The construct related evidence will be studied in line with the dimensionality of the items, traits measured by the items and finally the stability of these dimensions across the years and gender groups. In this respect the test scores will be cross validated with reference to time and different groups of students taking the tests.

In the analysis, the first and the second section of the mathematics sub-tests in the year of 2006 were considered. The main objective of the study is to assess the content specifications via exploring the congruence between the SST content and mathematical skills that are theoretically defined in the literature. Moreover, as a secondary analysis, empirical evidence is aimed to collect for supporting whatever the dimensions being assessed by the test items. In this analysis, it is aimed to describe different sub-domains of mathematical skills considered in the SST. Moreover, the consistency of these skills across gender, school type and years was also considered for cross-validation purpose. Thus, the study focuses on the following questions;

1. What cognitive skills are assessed in the mathematics subtest of the university entrance examination?

2. What are the dimensions of mathematics subtests (first and second section) in the student selection tests?

3. Do dimensions of 2006 SST mathematics subtests provide stable factorial structure across gender groups and school types?

4. Do dimensions of 2006 SST mathematics subtests provide stable factorial structure across years?

5. What mathematical skills are achieved at different ability levels of the students?

6. What are the relationships among the dimensions defined in the mathematics subtests of the SST?

7. Do the structural relationships defined in the mathematics subtests hold across gender and school type?

1.2 Definition of Terms

All the specific terms that are used in following chapters are defined in details in this section.

The Student Selection Test (SST): SST is administered since 1974 by Student Selection and Placement Center. SST is a paper and pencil test in which students mark their answers to optically readable answer sheets. There is a time limit for the whole test, however total time can be used for any subtests. The test is administered once a year and for each year test is constructed with new items. In 2006 the structure of SST has changed. The new exam is composed of 8 sub-tests, each includes 30 questions. These subtests are Turkish Language, Social Sciences-I, Mathematics-I, Natural Sciences-I which are related to basic common courses (first section) and Literature-Social Sciences, Social Sciences-II, Mathematics-II, Natural Sciences-II which are related to advanced subject-area courses (second section). However, for a student, it is necessary to answer 6 of these sub-tests, therefore there are total of 180 questions for each students. A student who wants to have a Science Score will answer all four basic courses sub-tests from first section, Mathematics-II and Natural Sciences-II subtests from second section; a student who wants to have a Turkish-Mathematics score will answer all four basic courses sub-tests from first section, Literature-Social Sciences and Mathematics-II subtests from second section; a student who wants to have a Social score will answer all four basic courses sub-tests from first section, Literature-Social Sciences and Social Sciences-II subtests from second section. Therefore, both two mathematics sections (Mathematics-I and Mathematics-II) are answered only by students who want to have Science score or Turkish-Mathematics score.

Mathematics Achievement: Mathematics achievement is measured by students' answers to Mathematics-I and Mathematics-II sections of SST.

School Type: Private and Public schools are identified according to Table 6, School Types and codes, which is published by Student Selection and Placement Center (Student Selection and Placement Center, 2006d).

Poor Items: Items which have corrected-item total correlation value lower than 0.200 are defined as poor items.

1.3 Significance of the Study

The main purpose of this study is to evaluate the content of the mathematics subtest of the SST in line with mathematical cognitive skills and eventually provide construct related evidence for the test scores. This thesis first that investigates what is measured by newly structured SST mathematics sections and analyzing construct validity of SST Mathematics sections by proposing and testing several nested confirmatory factor models.

Identifying what is measured by SST mathematics sections is important in the aspect that the institution responsible for the exam and item developers will have opportunity to check whether what they aimed to measure is achieved or not. Especially, identifying factors that are not measured properly, if any, will be an important feedback for institution and policymakers. These feedbacks will be helpful for revising future test plans and item writing procedure.

Tests like SST have to behave equally to different subgroups. Cross-validation analysis of this study that investigates similarity of what is measured across years and across groups will strengthen importance and significance of results.

CHAPTER 2

REVIEW OF RELATED LITERATURE

Four main sections were included in this chapter. These sections are mathematical abilities in standard tests; studies about the Student Selection Test (SST); studies about construct validity of tests; studies about invariance across groups and group differences.

2.1 Mathematical Abilities in Standard Tests

In the literature, mathematical skills are basically considered within problem solving processes. Krulik and Rudnick (1989) defined problem solving as the process in which previously gained knowledge, understanding and skills are used to compete with new unfamiliar situation. They added that in this process students should successfully synthesize their learning and apply it to new challenges. Noddings (1985) stated that purpose of problem solving is not only reaching to the solution of a problem, therefore, in classroom teaching, more emphasis should be given to the *process* of problem solving. Rubinstein (1980) underlined that problem solving ability is acquired if a student can transfer and apply school learning to real life situations and problems. Similarly Schwieger (1999) stated that educators discover the importance of preparing students to cope with real life problems related to mathematics. It is realized that solving in class mathematics questions does not guarantee solving real life mathematics problems for students.

In standardized tests, assessment of problem solving abilities takes important part. For instance, in PISA 2003, important emphasis is given to problem solving skills domain. It is explained that language, mathematics and science are important foundations of knowledge and skills for school subjects however, much wider types of competencies is needed for students to be successful in future challenge. One of the most important competencies of future life challenge is problem solving skills. Problem solving is seen as “providing an essential basis for future learning, for effectively participating in society and for conducting personal activities”. Types of problems defined in PISA 2003 are decision making, system analysis and design, and trouble shooting. In order to give feedback, three proficiency levels described for problem solving skills. Result of PISA 2003 showed that, more than one third of 15-year-old students in some countries are in high level of problem solving, however, in some countries majority of students even can not be classified as basic problem solvers. Half of the students participating from Turkey can not reach even level 1, where level 3 is described as highest level for problem solver. An example of problem solving item used in PISA 2003 is given in Appendix E (OECD, 2004).

In Cito Türkiye Pupil Monitoring System (ÖİS), problems solving is one of the important dimension of mathematics sub-domains like numbers, measurement, geometry, and probability & statistics. It is explained that in ÖİS there are three indispensible properties of problems: problems should be related to daily life situations; problems should be meaningful and concrete to students; a decision making process should be a part of problem solving process. A situation other than above description can not go beyond practice of algorithmic calculation. It is also stated that, generally, in curriculum, emphasis is given to steps of problem solving rather than correct definition of it. Therefore, items related problem solving and algorithmic calculation are mixed each other in some standardized tests of Turkey (İş Güzel, 2009).

In this section, also, mathematical abilities measured in many standardized tests are described. These mathematical abilities constitute assessment framework for many standardized tests. Similarities and differences related to mathematical abilities between these standard tests and SST will be discussed in chapter five by using empirical results.

In PISA 2003, mathematics assessment, like other sub-domains, focus on determining whether students can use their learning in the daily life situations they are likely to encounter. This can be situations people face when they shopping, travelling, cooking, dealing with their personal finances, judging political issues, etc. in which the use of quantitative or spatial reasoning or other mathematical competencies to clarify, formulate or solve problems. In mathematics assessment framework of PISA 2003, what is measured in mathematics section is explained. Three dimensions is defined which are the *situations or contexts* in which the problems are located; the *mathematical content* that has to be used to solve the problems; the *competencies* that have to be activated in order to connect real world, in which the problems are generated. With respect to mathematical competencies, cognitive processes that are measured are described as “competency clusters” which are the reproduction cluster, the connections cluster, and the reflection cluster. Reproduction cluster includes standard representations and definitions, routine computations, routine procedures, routine solving problem. Connections cluster includes modeling, standard problem solving translation and interpretation, multiple well-defined methods. Reflection cluster includes complex problem solving and posing, reflection and insight, original mathematical approach, multiple complex methods and generalization (OECD, 2003).

In TIMSS 2007, cognitive skills that are measured in mathematics classified as three main domains that are *knowing*, which covers the facts, procedures, and concepts students need to know; *applying*, which focuses on the ability of students to apply knowledge and conceptual understanding to solve problems or answer questions; *reasoning*, which goes beyond the solution of routine problems to encompass unfamiliar situations, complex contexts, and multi-step problems (Mullis et al. 2007).

Bloom et al (1971) and TIMSS 2007 have close definitions about cognitive processes. Bloom et al (1971) stated that cognitive processes Knowledge of specific facts, Knowledge of terminology, Ability to carry out algorithms is called "Computation/Knowing"; Knowledge of concepts, Knowledge of principles, rules and generalizations, Knowledge of mathematical structure, Ability to transform problem elements from one mode to another, Ability to follow a line of reasoning, Ability to read and interpret a problem is called "Comprehension/Knowing"; Ability to solve routine problems, Ability to make comparisons, Ability to analyze data, Ability to recognize patterns isomorphisms, and symmetries is called "Application/Applying"; and Ability to solve nonroutine problems, Ability to discover relationships, Ability to construct proofs, Ability to criticize proofs, Ability to formulate and validate generalizations is called "Analysis/Reasoning".

Student Selection and Placement Center stated that in SST, first section of mathematics has items about "power of using mathematical relations" and second section of mathematics has items about "Mathematics and Geometry" (Student Selection and Placement Center, 2006e).

2.2 Studies about the Student Selection Test

Berberoğlu (1995) and Berberoğlu et al. (1996) studied SST 1992 Mathematics subtest for subgroups. In the gender comparison, it is found that males are better at computation items whereas females are better at word problems and geometry items. In the SES comparison, it is found that high SES groups are better at most of the items, especially the word problems. It is also stated that Mathematics items in SST 1992 could be categorized only in comprehension level, none of the items could be categorized to measure application or analysis level defined by Bloom.

Uygun (2008) searched content-related and construct-related validity evidence for science subtests of 2006 SST. It is found that science items are generally measuring more than one content and more than one single cognitive process per item. It is also found that, what actually these items were measuring very hard to interpret because their cognitive processes were very close. Factor analysis results showed that items clustered according to their difficulty level. Besides that, significance mean differences across school types were obtained. Finally, it is stated that 2006 SST science section had a high internal consistency value of 0.94.

Kılıç (1999) conducted a study to investigate fit of one, two and three parameter models of item response theory to 1993 SST. It is concluded that items in SST is very difficult and not appropriate for ability level of students, especially mathematics and science items. For considering fit model analysis, the fit of the three parameter model found to be better for the subtests of SST. Can (2003) also investigated verbal section of 2001 SST with respect to IRT models. It is found that fit of one parameter model was better for verbal section of SST. Besides that, it is stated that verbal section items are moderately difficult for students.

Köksal (2002) investigated biology items of 1998 SST – 2001 SST with respect to cognitive processes and subject matters. Besides that, gender performance difference across cognitive processes is also investigated. It is found that items that are named to be measuring higher order thinking skills and items that are named to be measuring lower level cognitive process by experts were loaded to same component. It is stated that although there are some difference between scores on different subdomains, males and females are generally not successful in science.

2.3 Studies about Construct Validity

There are very limited studies about construct validity of SST. Berberoglu (1996) studied SST in terms of technical characteristics with main emphasis on the construct validity and gender bias issues. It is stated that mathematics section of SST was measuring "ability to make use of basic mathematical concepts and rules". It is also stated that items in mathematics subtest can be generally categorized as computation, word problems and geometry items (17 items, 5 items and 10 items respectively). Factor analysis results in order to assess cognitive characteristics measured by SST showed that computation and word problem items were generally loaded on a common factor, however, geometry items were loaded on a separate factor. It is concluded that SST measures a multidimensional trait and SST has content-wise organization.

Aslan (2000) conducted a study in which the construct validity of 1998 SST is investigated through results of exploratory factor analysis. Besides that, cross validation of results across gender is investigated. It is found that Turkish and social science items loaded to one factor; mathematics and science items are loaded to three different factors. It is stated that difficulty of items was interfere to the loadings. It is also found that factorial structure of SST was not different between gender groups.

Tuna (1995) searched empirical evidence for the cognitive characteristics underlying 1993 SST scores of students. Results of exploratory factor analysis showed that factors can be classified by item difficulty, content and taxonomic levels. It is important that when results of exploratory factor analysis investigated for mathematics items, computation and word problems generally grouped on a common factor whereas geometry items created a different factor. Also, no significant difference between the item test correlations across gender is found. Besides that, the structure of test content which is assessed by reliability estimates was equal across gender.

Several studies that investigate construct validity of other measurement instruments are also presented. Carlstedt and Gustafsson (2005) conducted a study related to construct validation of the Swedish scholastic aptitude test (SweSAT). It is stated that, although SweSAT is an important examination because of its use to pass higher education, SweSAT lacks a theroretical basis for its construction. Besides that, there is not much information about what it actually measures. They proposed four types of models to study construct validity of SweSAT. It is found that general ability and crystallized intelligence are strongly represented in the SweSAT, however, third proposed dimension which is general visualization is not.

Chen and Thompson (2004) presented a paper related to examining the construct validity of scores on self-concept scale for elementary students. They carried out confirmatory factor analysis for three alternative factor models. It is concluded that three-factor oblique model fits the data better than other models. Also, factorial invariance analyses across gender and grade groups are conducted. Fit indices showed that factor loadings, factor variances and error variances across gender groups and grade groups were invariant.

2.4 Studies about Invariance across Groups

Gender difference studies are one of the most important parts of comparative group research. Some selected studies related gender differences on standardized tests and academic performances are summarized below. Mau and Lynn (2001) reported that in late adolescence and early adulthood, males gets higher scores than females on college aptitude tests of SAT and ACT, which consists of general cognitive and reasoning ability. However, it is also reported that females gets higher grades during college education. Therefore, it is stated that if an assessment instrument is related cognitive tests, males will probably have an advantage and if it is related to coursework, females will probably have an advantage. Brynes and Takahari (1993) stated that the females have the advantages on tests that require computational skills whereas males have the advantages on tests that require problem solving.

Dayıoğlu and Aşık (2004) studied academic performance of gender groups in Turkey. It is reported that males had higher university entrance scores between 1996 and 2002. However, consistent with the literature, it is stated that females have higher Cumulative Grade Point Average than males in undergraduate programs of METU.

From 1997 to 2004, mean SAT-Math scores of males was higher than females. Average SAT-Math scores of males were ranging from 530 to 537, whereas average SAT-Math scores of females were ranging from 494 to 503. Minimum difference between these scores was 34 (World Almanac & Book of Facts, 2005). For the class of 2006, it is reported that average SAT-Math score for males was 536, whereas it was 502 for female counterparts (College Board, 2006).

Stricker et al. (2005) studied the factor structure of LanguEdge test and invariance of its factors across language groups. It is stated that issues like whether four section of this test measuring different constructs and whether the same construct is assessed in different language groups have not been addressed until this study. It is concluded by the confirmatory factor analysis, four section of LanguEdge test represent two correlated factors, which are speaking and a fusion of Listening, Reading and Writing. Besides that, the number of factors, the factor loadings and the error variances were invariant across groups whereas factor correlations were not invariant.

Stricker and Rock (2008) studied the factor structure of TOEFLiBT for invariance of its factors across language groups. Five different confirmatory factor models are tested and whether the factor structure is invariant across different language groups is investigated. It is concluded by confirmatory factor analysis, the number of factors, the factor loadings and the error variances and factor correlations were invariant across groups.

Kollu (2006) investigated effects of private schools on achievement of public schools in Turkey. Achievement in Kollu's study is defined by scores on 2003 SST. It is claimed that private schools have negative effect on public school achievement. Besides that, according to Çınar (2006) students in public schools have opinion that the educational services of the public schools insufficient to become successful in the SST.

Thus the validity of mathematics subtest of the SST will base on content-wise evaluation of the test item content with respect to mathematical cognitive skills and abilities emphasized in the literature; and base on several confirmatory factor analysis models that are tested by using relevant studies and theories in the literature.

CHAPTER 3

METHODOLOGY

This chapter presents the methodology of the study. Sections in this chapter are population and sample, instruments, and procedures.

3.1 Population and Sample

This study examined SST mathematics sections data for 2006 SST and 2007 SST separately. SST is administered to students graduated from high school. Detailed information about sample of this study is given below.

3.1.1 Sample in 2006 SST

Total of 1 511 596 students took 2006 SST. The students who were graduated from Turkish-Mathematics and Science branches in the secondary school and responsible of answering questions in both mathematics sub-tests constitute the sample of the study. Therefore, 2006 SST data set in this study has 872 956 subjects. Among 872 956 subjects, 54.6% of them is male and 45.4% of them is female; 14.5% of them is private school students and 85.5% of them is public school students.

3.1.2 Sample in 2007 SST

Total of 1 614 406 students took 2007 SST. The students who were graduated from Turkish-Mathematics and Science branches in the secondary school and responsible of answering question in both mathematics sub-tests constitute the cross validation sample of the study. Therefore, 2007 SST data set in this study has 915 161 subjects. Table 3.1 shows demographic information for subgroups in 2006 SST data

Table 3.1 Demographics of the subgroups in 2006 SST

		N	Ratio
	Total	872 956	
Gender	Males	476 432	54.6 %
	Females	396 524	45.4 %
School Type	Private	126 765	14.5 %
	Public	746 191	85. 5%

3.2 Instruments

This study examined mathematics items in 2006 SST mathematics sections and in 2007 SST mathematics sections. Detailed explanations about structure of these tests and sections are given below.

All mathematics items in 2006 administration and in 2007 administration are given in Appendix A through Appendix D. For the convenience of the reader, the following representations for the items are used throughout the thesis: items in first section are indicated as MT1.x and items in second section are shown as MT2.x where x stands for order of item in that section.

2006 SST and 2007 SST mathematics exams consist of two sections which are called Mathematics-I and Mathematics-II. Both sections consist of 30 items, therefore there are total of 60 mathematics items in the tests. Mathematics-I consists of items related to basic mathematics topics which are covered by all students graduated from high schools. Mathematics-II consists of items related to advanced mathematics topics which are covered by students who graduated from Science or Turkish-Mathematics branches of high schools.

There is very limited information about measurement properties of SST. Cognitive processes dimensions behind item groups and their relation to content dimension is not clear. Student Selection and Placement Center only stated that in SST, first section of mathematics has items about "power of using mathematical relations" and second section of mathematics has items about "Mathematics and Geometry" (Student Selection and Placement Center, 2006e).

3.3 Procedure

In this section methods used for each research question are presented.

R1: The first research question asks "What cognitive skills are assessed in the mathematics subtest of the university entrance examination?". The analysis for this question will base on researcher's and experts' content-wise evaluation of the test item content with respect to mathematical cognitive skills emphasized in the literature.

R2: The second research question asks "What are the dimensions of mathematics subtests (first and second section) in the student selection tests?". The analysis for this question will base on examining underlying constructs of mathematics sections by the factor analysis. In order to perform this analysis, first of all, exploratory factor analysis will be conducted. By using results of exploratory factor analysis, items related to each definable constructs will be grouped according to cognitive processes dimensions and content dimensions. Then, several confirmatory factor analysis models will be tested by using relevant theory. After that, these nested models will be compared with each other to find out the number of dimensions in SST mathematics sections. A model with best fit index values will be used for further interpretations of the sub-dimensions of the SST. Finally, dimensions of mathematics subtests will be identified.

R3: The third research question asks "Do dimensions of 2006 SST mathematics subtests provide stable factorial structure across gender groups and school types?". The analysis for this question will base on performing the test of whether the same factor analysis model holds for gender groups and school types in terms of number of factors, factor correlations, error variances, and factor loadings. By using results of equality of factor structure analysis, differences or similarities across groups will be identified.

R4: The fourth research question asks "Do dimensions of 2006 SST mathematics subtests provide stable factorial structure across years?". The analysis for this question will base on conducting confirmatory factor analysis for 2007 SST mathematics sections. By using results of confirmatory factor analysis, similarities and differences across 2006 SST and 2007 SST data will be investigated.

R5: The fifth research question asks "What mathematical skills are achieved at different ability levels of the students?". The analysis for this question will base on estimation of 2006 SST item parameters by using Item Response Theory (IRT). The item mapping procedure with P50 and P80 response probabilities will be used for this analysis. The aim is to describe what mathematical skills are achieved at different levels of the test scores-ability levels. The analysis will provide a base to infer about the prerequisite nature of the mathematical skills.

R6: The sixth research question asks "What are the relationships among the dimensions defined in the mathematics subtests of the SST?". The analysis for this question will base on testing a structural equation model about proposed relations among dimensions in the mathematics subtests.

R7: The seventh research question asks "Do the structural relationships defined in the mathematics subtest hold across gender and school type?". The analysis for this question will base on testing the equality of the structural equation model across gender groups and school types.

3.3.1 Descriptive Summary

As a descriptive summary, number of student, minimum scores, maximum scores, mean scores, mean scores per item, standard deviation of scores, skewness and kurtosis values of item groups are presented.

3.3.2 Reliability Analysis

Several methods can be used to measure reliability of an instrument. Cronbach's alpha (or Coefficient alpha) is one of the important indicators of reliability which is designed as a measure of internal consistency. In this measure the question of whether all the items within the instrument measure the same construct or trait is asked. As Cronbach's alpha closes to the value of 1.00, the higher the internal consistency of items in the instrument (George &Mallery, 2001).

In standard tests, subtests should be composed of items which have high correlation with the rest of the items. DeVellis (2003) stated that the corrected item-total scale correlation correlates the item and the total score with excluding that item from total score. Pallant (2007) stated that corrected item-total correlation values shows the degree an item correlates with the total score in which a low value shows that this item measures something different from the total test. In their study, Butler et al. (2008) excluded items with corrected-item total correlations lower than 0.200 in order to have better scale scores.

In this study, items that have corrected-item total correlations lower than 0.200 are eliminated to see effects of poor items to overall reliability and to fit of confirmatory factor model.

3.3.3 Summated Scales

Summated scales for a student can be obtained by adding number of scores for a predefined group of items. There are several advantages of using summated scales. Firstly, by adding item-scores, a larger differentiation and variation in the measurement is obtained. Secondly, according to the central limit theorem, as the number of variable increases, the sum of several variables will probably approach to a normal distribution and obtained new scores will be interval scale scores (Blunch, 2008).

SSPC uses correction for guessing when they calculate raw scores of students. Raw student scores are calculated by recoding correct answers to “1”, wrong answers to “-0.25”, and missing answers to “0”. This recoding and correction affects students’ behavior when they give answers to items, because, even very minor score differences affects whether a student admitted or not. With this fact, students tend to omit an item if they are not sure about truth of their choice. Therefore, a wrong answer and a missing answer do not have same meaning on SST. For this reason, as SSPC does, in this study correct answers are coded as “1”, wrong answers are coded as “-0.25”, and missing answers are coded as “0”.

3.3.4 Confirmatory Factor Analysis

Confirmatory Factor Analysis (CFA) is used to test nested competitive models and confirm a theory or potential relationships among variables. The main question in CFA is how well the collected data fit the hypothesized model. In other words, possibility of empirically confirmation of the hypothesized model is searched (Sharma, 1996). Kline (1998) stated that a model is established at the beginning of the analysis and main purpose is to test whether this model is supported by the data.

Jöreskog and Sörbom (1993) stated that there are three general applications of CFA; (1) strictly confirmatory, (2) alternative models, (3) model-generating. The first situation occurs when the researcher has only one model that is accepted or rejected based on data. The second situation occurs when more than one model is available. The last situation occurs when the proposed model does not fit the data and is modified by the researcher. This new model is tested again using same data.

In CFA, there are parameters in the hypothesized model to be calculated, namely, the regression coefficients (factor loadings), the variances and covariances of independent variables. These parameters are estimated using sample data to represent best possible population values. Then, these estimated parameters are used to produce an estimated population covariance matrix $\sum$. After that, this population covariance matrix $\sum$ is compared with the sample covariance matrix S and if the difference is small and not statistically significant, the model is validated (Ullman, 2001). Important practical issues related confirmatory factor analysis and structural equation modeling (SEM) is discussed below.

3.3.4.1 Assumptions of Confirmatory Factor Analysis

In order to have valid CFA and SEM results, multivariate normality, and linearity assumption should hold. Observed variables should be continuous and interval scaled. Ratio of sample size to number of variable is also important.

Multivariate normality is assumed by many of the estimation techniques used in CFA (Ullman, 2001). Multivariate normality means that each observed variable and all linear combinations of these variables should be normally distributed. However, it is impractical to test an infinite number of linear combinations of variables for normality (Tabachnick & Fidell, 2001). In order to assess normality of each observed variables individually, either statistical or graphical methods can be used. George and Mallery (2001) stated that skewness and kurtosis value between ± 1 is considered as excellent and between ± 2 is considered acceptable for normality. Also, it is important to note that when sample size is very large, the impact of departure from zero skewness and kurtosis does not affect the results of analyses (Tabachnick & Fidell, 2001).

CFA techniques uses linear relationships among variables therefore, linear relationships among pairs of observed variables should be assessed by using scatterplots. CFA assumes that observed variables are continuous and they measured on an interval scale. CFA analyzes results are less stable when sample size is small. Fewer than 10 subjects per estimated parameter may be adequate if the observed variables are normally distributed (Ullman, 2001). MacCallum et al. (1996) provided table of minimum required sample sizes for conducting CFA study in their article. For example, minimum sample size to achieve power of 0.80 for 50 degrees of freedom is 268.

3.3.4.2 Steps of Confirmatory Factor Analysis

1) Model Specification

The first and very important step in CFA is model specification. By using relevant theory, research and available information, a specific model that will be tested or confirmed has to be specified. Therefore, in this step it is decided which variables are included in the analysis and how they are related to each other by specifying relationships. One important possible problem in this step is specification error which occurs when an unimportant variable is included or an important variable is excluded from the model. A misspecified model will produce biased parameter estimates which will be systematically different from actual values in true model. If a model is misspecified, it is most probably that this model will not fit the data (Schumacker & Lomax, 2004).

2) Model Identification

A model is identified if there is a unique solution for each parameter in the model on the basis of the sample data which produce the sample covariance matrix S and the theoretical model implied by the population covariance matrix $\sum$. One necessary condition of getting a unique solution is having more data points than number of parameters that are estimated. In other words model should be "overidentified". The number of data points is the number of sample variances and covariances. The number of parameters that are estimated is sum of number of regression coefficients, variances and covariances (Ullman, 2001; Schumacker & Lomax, 2004). If the number of data points is equal to the number of parameters that are estimated, the model is called "just identified" in which, adequacy of model can not be tested. In this case, only hypotheses about certain paths in the model can be tested. If the number of data points is less than number of parameters that are estimated, the model is called "underidentified" in which, parameters can not be estimated. In this case, one possible solution to make a model overidentified is to set a parameter to a specific value or to constrain value of a parameter equal to value of another parameter (Ullman, 2001).

3) Model Estimation

In these step, parameters are estimated using sample data and specified model. It is desired to get estimates that produce implied covariance matrix $\sum$ which is very close to sample covariance matrix. Minimum difference between elements in the matrix S and the elements in the matrix $\sum$ is desired. Several estimation methods are available (Schumacker & Lomax, 2004). Maximum Likelihood (ML) and Generalized Least Squares (GLS) are mostly used estimation methods in SEM. Sample size and the normality assumption are important when selecting appropriate estimation method. If the normality assumption is not violated, ML performed well over sample size 500, and GLS performed better when the sample size is less than 500. However, if the normality assumption is violated ML and GLS work well when sample size is more than 2500 but GLS is slightly better with smaller sample sizes (Ullman, 2001).

4) Model Testing

After a model is specified, identified and parameters are estimated, it is important to assess whether this model is a “good” one. Good model means that difference between sample covariance matrix and the population covariance matrix is minimum. This means that there is a fit between these matrices. Therefore, obtained sample data fit the theoretical model (Ullman, 2001).

There are two dimensions of model fit. The first one is globally testing fit of entire model. The second one is individually testing fit of each parameter in the model. There are many fit indices to test first one and these indices are explained below in detail. In testing individual parameters, it is expected that each individual parameter is significantly different from zero, sign of the parameter is in expected direction, parameter estimates are within an expected range of values like variances should not have negative values and correlations should have values between -1 and 1 (Schumacker & Lomax, 2004).

As it is stated, the goal in CFA is to construct a model that fits the sample data. Therefore, minimum difference between sample covariance matrix S and population covariance matrix $\sum$, in other words, a nonsignificant chi-square is desired. However, chi-square values are highly inflated when the sample size is large. Blunch (2008) stated that if the sample size is very small, any model will be accepted and if the sample size is very large, any model will be rejected. For this problem, lots of fit indices are developed that examine model fit while eliminating or minimizing the effect of sample size. One indicator of good fitting model is when the ratio of the chi-square value to the degrees of freedom is less than 2 (Ullman, 2001).

The independence chi-square test value should be always significant. Null hypothesis in this test is that there is no relationship among variables. Therefore, significant independence chi-square test means that there is some relationship among variables (Ullman, 2001).

The goodness-of-fit index (GFI) and the adjusted goodness-of-fit index (AGFI) are other important fit indices. They account for proportion of variance in the sample covariance by the estimated population covariance matrix. These indices can be considered as analogous to R^2 in multiple regression. AGFI is adjusted version of GFI for the number of parameter estimated (Ullman, 2001). GFI values of 0.90 and more and AGFI values of 0.80 and more means good fitting model (Segars & Grover, 1993). Schumacker and Lomax (2004) stated that GFI values of 0.95 and above is a sign of good model fit. AGFI values higher than 0.90 are acceptable for good fit (McDonald & Moon-Ho, 2002).

The root mean square residual (RMR) is a residual-based fit index in which the average difference between the sample variances and covariances and the estimated population variances and covariances is calculated (Ullman, 2001). RMR values of 0.10 and less means good fitting model (Segars & Grover, 1993). Blunch (2008) and Byrne (1998) stated that RMR values less than 0.05 is a sign of good fit. The standardized root mean square residual (SRMR) value of 0.08 and lower is a sign of good fit (Hu and Bentler, 1999).

The root mean square error of approximation (RMSEA) estimates the lack of fit by comparing perfect (saturated) model and estimated model using degrees of freedom. RMSEA ranges from 0 to 1 and values less than 0.06 means a good fitting model (Ullman, 2001). If RMSEA value is higher than 0.10, this means a poor fitting model (Browne and Cudeck, 1993).

Incremental fit index (IFI) and comparative fit index (CFI) are mostly advised comparative fit indices in which both of them includes comparison of independence model and estimated model with using degrees of freedom. Both of them range from 0 to 1 and values over 0.95 means good fitting model (Ullman, 2001). Normed fit Index (NFI) and Relative fit index (RFI) is also used to compare a restricted model with full model (Schumacker & Lomax, 2004). NFI and RFI values larger than 0.95 also sign of a good fit (Blunch, 2008)

Another fit index is Expected Cross-Validation Index (ECVI). ECVI is used to assess the likelihood that a proposed model in a single sample will cross-validates with same population of close sample size. To evaluate ECVI values, ECVI index is calculated for several models and a model with the smallest ECVI value has the greatest possibility to cross-validate. Therefore, smallest value of ECVI is better (Byrne, 1998).

Schumacker and Lomax (2004) gave the formulas for most of these fit indices with using χ^2 of the null model (covariances are assumed to be zero in the model), *df* of null model, χ^2 of hypothesized model, *df* of hypothesized model and sample size.

$$\text{GFI} = 1 - (\chi^2_{model} / \chi^2_{null})$$

$$\text{NFI} = (\chi^2_{null} - \chi^2_{model}) / \chi^2_{null}$$

$$\text{RFI} = 1 - [(\chi^2_{model} / df_{model}) / (\chi^2_{null} / df_{null})]$$

$$\text{IFI} = (\chi^2_{null} - \chi^2_{model}) / (\chi^2_{null} - df_{model})$$

$$\text{CFI} = 1 - [(\chi^2_{model} - df_{model}) / (\chi^2_{null} - df_{null})$$

$$\text{RMSEA} = (\sqrt{[\chi^2_{model} - df_{model}] / [(N-1) df_{model}]}$$

5) Model Modification

When the fit of the implied model is not satisfactory according to several fit indices, then it is necessary to modify the model in order to improve the fit. In the process of modifying a model, the change that is made should make sense, in other words, it should be supported by theory and there should be an explanation. Otherwise, it might be hard to make a conclusion about that relationship. One method of modification is to eliminating nonsignificant relations. However, it is important to note that, if a parameter is not significant but important according to theory, it should be kept in the model. Another method is examining standardized residual matrix, which is the standardized difference between the observed covariance matrix *S* and the implied covariance matrix $\sum$. Large standardized residuals (larger than 1.96 or 2.58) show some problems. The well-known another procedure for model modification is to use modification indices which are produced by CFA softwares. These indices show effects of expected change in the fit when a specified change related a parameter is made (Schumacker & Lomax, 2004).

3.3.4.3 Multiple Group Analysis

The main hypothesis tested by multiple group analysis in CFA is whether the data from different groups comes from the same population. In other words, whether different groups have same measurement model is tested. After selecting the factor model that is best supported by the data, this model will be tested about the invariance of the factors across different samples. In multiple group analysis, it is wanted to investigate to what extent (1) number of factor is invariant; (2) the correlations between latent variables are invariant; (3) the error variances of the observed variables are invariant across groups; (4) the factor loadings corresponding to the paths from latent variables to the observed variables are invariant.

In this study, the purpose of conducting multiple group analysis is to assess the invariance of the defined and confirmed factors in SST Mathematics sections for subgroups defined by a) gender b) school type.

In order to test invariance of factor structures the following models will be tested for each subgroup.

Model A: The number of factors is invariant,
Model B: The factor correlations are invariant,
Model C: The factor correlations and error variances are invariant,
Model D: The factor correlation, error variances and factor loadings are invariant.

CHAPTER 4

RESULTS

This chapter is divided into seven main sections related to research questions. The first section is related identifying cognitive skills assessed in mathematics subtests of SST. In this section content-wise evaluation of researcher and experts with respect to mathematical cognitive skills is presented. The second section is related identifying dimensions of mathematics subtests empirically. As a results of exploratory factor analysis items are grouped, and confirmatory factor analysis is conducted to further confirm the existing dimensionality of the test items. The third section is related cross validating constructs measured in 2006 SST across groups. In this section multiple group analysis based on accepted measurement model for gender groups and school types is conducted. The fourth section is related cross validating constructs measured in 2006 SST across years. In this section confirmatory factor analysis of accepted measurement model with using 2007 SST data is re-conducted. The fifth section is related investigating relations between mathematical skills and ability levels of students. In this section all mathematics items in 2006 SST are analyzed by using Item Response Theory (IRT). The item mapping procedure with P50 and P80 response probabilities is used for this analysis. The aim is to describe what mathematical skills are achieved at different levels of the test scores-ability levels. The analysis is a base to infer about the prerequisite nature of the mathematical skills. The sixth section is related identifying relations among dimensions defined in the mathematics subtests. In this section, structural equation modeling among mathematical dimensions according to prerequisite nature of the mathematical skills is tested. Finally, the seventh section is related cross validation of structural relationships across groups. In this section analysis related to equality of structural equation across gender groups and school type is investigated.

4.1 Cognitive Skills in 2006 SST

The first research question asks “What cognitive skills are assessed in the mathematics subtest of the university entrance examination?”. In order to answer this question content-wise evaluation of the test items with respect to mathematical cognitive skills emphasized in the literature is used. For this evaluation researcher’s and experts’ opinions are taken considering the cognitive skill classifications of Bloom et al. (1971) and TIMSS 2007. OSYM prepares items with respect to the subject areas of the secondary school courses and Bloom’s taxonomy of educational objectives (Berberoglu, 1996). Therefore, even though there are other frameworks, in this study, Bloom’s taxonomy of educational objectives is preferred for this evaluation.

Bloom et al (1971) and TIMSS 2007 classified cognitive skills in Mathematics assessment in four main categories which are “Computation/Knowing”, “Comprehension/Knowing”, “Application/Applying”, and “Analysis/Reasoning”. Knowledge of specific facts, Knowledge of terminology, Ability to carry out algorithms are classified as “Computation/Knowing”; Knowledge of concepts, Knowledge of principles, rules and generalizations, Knowledge of mathematical structure, Ability to transform problem elements from one mode to another, Ability to follow a line of reasoning, Ability to read and interpret a problem which are classified as “Comprehension/Knowing”; Ability to solve routine problems, Ability to make comparisons, Ability to analyze data, Ability to recognize patterns isomorphisms, and symmetries which are classified as “Application/Applying”; and Ability to solve nonroutine problems, Ability to discover relationships, Ability to construct proofs, Ability to criticize proofs, Ability to formulate and validate generalizations which are classified as "Analysis/Reasoning".

In this first research question it is aimed to evaluate what extent the items in the SST mathematics sections are measuring these cognitive skills. When item content measured in SST compared with cognitive skills emphasized in Bloom et al (1971) and TIMSS 2007, it is evaluated that there are many items that measures ability to carry out algorithm. Bloom et al. (1971) stated that ability to carry out algorithm is most important subcategory of "Computation/Knowing" cognitive process. Solving linear equations, fraction operations, numerical calculations are some examples that represent ability to carry out algorithm. It is claimed that items similar to item MT1.2 are examples for this cognitive process.

Sample item for "Computation/Knowing" cognitive process

MT1.2

$$\frac{1-\frac{1}{2}+\frac{1}{3}}{\frac{1}{2}-\frac{1}{3}+\frac{1}{4}}$$

işleminin sonucu kaçtır?

A) 2 B) 1 C) 0 D) −1 E) −2

In SST, also, there are several items which measures knowledge of principles, rules, and generalizations. Bloom et al. (1971) stated that knowledge of principles, rules, and generalizations is one of the most important subcategory of "Comprehension/Knowing" cognitive process. Identifying relationships among concepts and problem elements represent this ability. It is claimed that items similar to item MT2.5 are examples for this cognitive process. Besides that in SST, there are items which measures ability to solve routine problems. Bloom et al. (1971) stated that ability to solve problems is one of the most important subcategory of "Application/Applying" cognitive process. A problem solving process that is encountered during the course of instruction in which the student is asked to carry out an algorithm to reach a solution represents this ability. It is claimed that items similar to item MT1.17 are examples for this cognitive process.

Sample item for "Comprehension/Knowing" and "Application/Applying" cognitive process

MT2.5	MT1.17
A boş olmayan bir küme olmak üzere, A dan A ya f ve g fonksiyonları tanımlanmıştır. $(fog)(x) = f(g(x))$ **ile verilen fog bileşke fonksiyonu <u>bire bir</u> ise aşağıdakilerden hangisi <u>kesinlikle</u> doğrudur?** A) f örtendir. B) g örtendir. C) f bire birdir. D) g bire birdir. E) gof bire birdir.	Bir araç, iki kent arasındaki yolu saatte ortalama 60 km hızla gidip, hiç mola vermeden saatte ortalama 80 km hızla dönerek yolculuğu 7 saatte tamamlıyor. **Bu iki kent arasındaki uzaklık kaç km dir?** A) 240 B) 280 C) 300 D) 320 E) 360

As a result of these evaluations and comparisons, researcher claimed that items in mathematics subtest do not match properly with theoretical framework of mathematics assessment emphasized by Bloom et al. (1971) and TIMSS 2007. As Table 4.1 shows, out of 60 mathematics items, researcher classified 45 items as "Computation/Knowing" cognitive skill, 5 items as "Comprehension/Knowing" cognitive skill and 10 items as "Application/Analysis" cognitive skill. None of the item is classified as “Analysis/Reasoning” cognitive skill which is mainly related to ability to solve nonroutine problems. This shows that distribution of items related to cognitive skills are not proper because majority of items are related to “Computation/Knowing” cognitive skills, whereas none of them is related to “Analysis/Reasoning” cognitive skill. This investigation shows that in SST mathematics sections, higher order thinking skills are not measured properly.

Table 4.1 Item Classifications According to Cognitive Skills

Cognitive Skills	Items
Computation/Knowing	MT1.1, MT1.2, MT1.3, MT1.14, MT1.4, MT1.5, MT1.6, MT1.7, MT1.11, MT1.12, MT1.13, MT2.1, MT2.2, MT2.3, MT2.4, MT2.6, MT2.7, MT2.9, MT2.10, MT2.11, MT2.12, MT2.13, MT2.14, MT2.15, MT2.16, MT2.17, MT2.18, MT2.19, MT2.20, MT2.21, MT1.21, MT1.22, MT2.23, MT2.24, MT2.25, MT1.23, MT2.26, MT2.27, MT1.24, MT1.26, MT2.28, MT2.29, MT2.30, MT1.29, MT1.30
Comprehension/Knowing	MT1.9, MT1.10, MT1.20, MT2.5, MT2.8
Application/Analysis	MT1.8, MT1.15, MT1.16, MT1.17, MT1.18, MT1.19, MT1.25, MT1.27, MT1.28, MT2.22

Content-wise evaluation of the test items with respect to mathematical cognitive skills emphasized in the literature is a subjective process. Therefore, in order to have more reliable results and get more opinions, items are asked to be classified by three experts. All of the experts had undergraduate degree related to mathematics; had graduate degree on measurement and evaluation. Besides that, two of these experts have doctorate degree on measurement and evaluation and one of them is a candidate to have a doctorate degree. A guideline is prepared by researcher and given to experts about how to classify items. Content dimension and cognitive process dimension is explained in details related to Bloom et al (1971) and TIMSS 2007.

It is seen that out of 60 items, 6 items are put into same group by all experts; 43 items are placed into same group by two experts (one expert had different opinion); and 11 items are placed into three different groups by these three experts. Therefore, congruence between experts generally achieved for 49 items. Out of 49 mathematics items, 5 items are related to “Computation/Knowing” cognitive skill, 22 items are related to “Comprehension/Knowing” cognitive skill, 18 items are related to “Application/Analysis” cognitive skill, and 4 items are related to “Analysis/Reasoning” cognitive skill. Coefficient of concordance calculated by average Kendall’s tau is 0.507 which shows moderate relation. Table 4.2 shows number of items classified by experts for each cognitive process.

Table 4.2 Number of Item Classifications by Experts

	Computation /Knowing	Comprehension /Knowing	Application /Applying	Analysis /Reasoning
Expert1	6	47	7	0
Expert2	8	21	19	8
Expert3	7	5	37	11

The item classification of researcher and experts has similarities in which items are mainly classified as either “Computation/Knowing” or “Comprehension/Knowing” cognitive skill (Researcher 83%, Experts 55%). Also, limited number of item is classified as “Analysis/Reasoning” cognitive skill by researcher and experts (Researcher 0%, Experts 8%). Researcher and experts agree on that in SST mathematics sections, higher order thinking skills are not measured properly. However, main classification difference between researcher and experts is that researcher classified items mainly as “Computation/Knowing” cognitive skill whereas experts classified items mainly as “Comprehension/Knowing” cognitive skill. Bloom et al. (1971) stated that dividing line between computation and comprehension cognitive skill is artificial and vague which explains this classification difference.

It is concluded that there is not high level of congruence between researcher and experts, and among experts on item classifications. What cognitive skills are assessed in the mathematics subtest of the university entrance examination is not totally definable by content-wise evaluation. Also, cognitive skills assessed by items in mathematics subtest do not match with theoretical framework of mathematics assessment emphasized in the literature. Therefore, in order to understand how items are grouped, a different approach is required. Empirical support for what is measured by SST mathematics section is necessary. In order to perform these analyses, as a first step, exploratory factor analysis is conducted.

4.2 Dimensions of Mathematics Subtests in 2006 SST

As was stated before, the second research question asks "What are the dimensions of mathematics subtests (first and second section) in the student selection tests?". In order to answer this question and examine underlying constructs of mathematics sections, exploratory factor analysis is conducted; items are grouped; and confirmatory factor analysis is performed.

4.2.1 Exploratory Factor Analysis of 2006 SST

In order to understand how items are grouped and to provide empirical support for what is measured by SST, exploratory factor analysis is conducted. SPSS 16.0 is used to conduct exploratory factor analyses. Kaiser-Meyer-Olkin measure of sampling adequacy value of 0.987 (marvelous), and significant Bartlett's test of sphericity statistics indicated that conducting factor analysis is appropriate (George and Mallery, 2001). Principal component analysis with varimax rotation showed that there are five components which have eigenvalues larger than 1.00. These components accounted for 38% of the total variance in the data. Table 4.3 shows rotated factor loadings for 2006 SST items. In this table, loadings less than 0.25 are omitted.

Exploratory factor analysis results in Table 4.3 indicate that five main dimensions are measured by SST mathematics sections. In order to investigate these dimensions, items will be grouped and five-factor model will be tested by confirmatory factor analysis.

Table 4.3 Rotated Factor Loadings for 2006 SST

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
MT2.17	0.622			0.255	
MT2.10	0.582	0.313			
MT2.6	0.553				
MT2.15	0.546				
MT2.16	0.545			0.262	
MT2.19	0.529	0.273		0.384	
MT2.22	0.513	0.432			
MT2.3	0.506			0.322	
MT2.8	0.501			0.301	
MT2.25	0.492				
MT2.9	0.465			0.406	
MT2.29	0.443	0.398			
MT2.7	0.432				
MT2.28	0.406				
MT2.12	0.400			0.253	
MT1.30	0.386	0.373			
MT2.14	0.369				
MT2.13	0.357				
MT1.11	0.324				
MT2.18	0.300				
MT2.20	0.262				
MT1.28		0.613			
MT1.29		0.572			
MT1.23		0.557			
MT1.25	0.324	0.542			
MT1.27	0.277	0.534			
MT1.26		0.517			
MT1.22		0.485		0.252	
MT2.30		0.468		0.367	
MT2.23	0.338	0.458			
MT1.19		0.435			0.302
MT2.27	0.319	0.423			
MT2.24	0.353	0.394			
MT1.7		0.306	0.747	0.266	0.304
MT1.3			0.704		
MT1.8			0.685		
MT1.13			0.674		
MT1.1			0.640		
MT1.2			0.606		
MT1.12			0.521	0.269	
MT1.9			0.501		
MT1.21		0.285	0.476		
MT1.18			0.414		0.260
MT1.24		0.333	0.347	0.357	
MT2.2	0.257			0.562	
MT2.21	0.404		0.391	0.528	
MT2.1			0.291	0.525	
MT1.14				0.510	
MT2.11	0.452		0.301	0.481	
MT2.26		0.403	0.308	0.421	
MT1.16		0.367	0.325	0.418	0.303
MT2.4				0.378	
MT1.6			0.314	0.375	
MT1.5		0.290		0.374	0.285
MT1.15		0.283			0.432
MT1.4			0.350		0.410
MT1.10			0.357		0.384
MT1.17		0.338		0.314	0.362
MT1.20	0.306	0.275			0.333
MT2.5					0.288

Loadings less than 0.25 are omitted

As was observed in the content analysis, the cognitive processes assessed by the items of SST do not match with the theoretical framework of mathematical skills. Thus, the researcher decided to group the items in a way to reflect content specification and the cognitive processes depicted in items. Total of thirteen item groups is established. Details about each item groups and process of grouping items according to results of factor analyses, what they measure, and the content is described below. Table 4.4 shows name of groups and items in each group.

While forming these groups and giving name to these groups, first content dimension and then cognitive process dimension is taken for reference. In content dimension, high school mathematics content before trigonometry topic is named as “basic”; items related to mathematical symbols is named as “symbolic”; high school mathematics content after trigonometry topic is named as “advanced”; and content of geometry is named as, “triangle”, “quadrangle”, “circle”, or “analytic geometry” according to related geometrical figures in item. In cognitive process dimension, classifications of Bloom et al (1971) and TIMSS 2007 is taken as a theoretical framework. Items related to ability to carry out algorithms, ability to carry out routine algebraic procedures are named as “calculations”; items related to knowledge of principles, rules and generalizations or ability to follow line of reasoning are named as “generalizations”; items related to ability to solve routine problems are named as “problems”. Further details are given below.

Group1: This group is named as **Basic Calculations**. Items MT1.1, MT1.2, MT1.3 and MT1.14 have common properties to be in this group. In all of these four items, a result of basic calculation using real numbers is asked. In order to find the result, several steps of calculations are necessary. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out algorithms can solve these items. Besides that, three of these items are loaded to factor 3 according to exploratory factor analysis results.

Group2: This group is named as **Symbolic Calculations I**. Items MT1.4, MT1.5, MT1.6, MT1.7, MT1.11, MT1.12 and MT1.13 have common properties to be in this group. In all of these items, result of a basic calculation using symbols, value of a symbol obtained by some calculations, or relation between symbols is asked. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out routine algebraic procedures can solve these items. Besides that, five of these items are loaded to factor 3 according to exploratory factor analysis results.

Group3: This group is named as **Basic Generalizations**. Items MT1.9, MT1.10 and MT1.20 have common properties to be in this group. In all of these items, a generalization by understanding defined situation is asked. The results obtained for these items are not pure calculation of some numbers or symbols. Similar to Bloom and TIMSS 2007 cognitive process definitions (Comprehension and Knowing), students who have knowledge of principles, rules and generalizations or ability to follow line of reasoning can solve these items. Besides that, two of these items are loaded to factor 5 according to exploratory factor analysis results.

Group4: This group is named as **Word Problems**. Items MT1.8, MT1.15, MT1.16, MT1.17, MT1.18 and MT1.19 have common properties to be in this group. In all of these items, a short passage which defines situation is given and by writing a mathematical expression related to this passage, performing steps of calculations is expected. Similar to Bloom and TIMSS 2007 cognitive process definitions (Application and Applying), students who have ability to solve routine problems can solve these items. Besides that, five of these items are loaded to factor 5 according to exploratory factor analysis results. Bloom et al (1971) stated that a problem that is encountered during the course of instruction in which the student is asked to carry out an algorithm to reach a solution is called routine problems. In non-routine problems, students are given a problem situation in which algorithmic solution is not available and requires the students transfer previous mathematics learning to a new context. Therefore, word problems in SST is considered as routine problems.

Group5: This group is named as **Symbolic Calculations II**. Items MT2.1 and MT2.2 have common properties to be in this group. In all of these items, result of a calculation done by using symbols is asked. In all of these questions in order to find the result, several steps of calculations are necessary. In this group different from Symbolic Calculation I, ability to conduct a series of factorization is measured. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out routine algebraic procedures can solve these items. Besides that, all of these items are loaded to factor 4 according to exploratory factor analysis results.

Group6: This group is named as **Advanced Calculations I**. Items MT2.3, MT2.4, MT2.6, MT2.7, MT2.9, MT2.10 and MT2.11 have common properties to be in this group. In all of these items, result of an advanced calculation (high school mathematics topics between trigonometry and limit of a function subject) is asked. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out routine algebraic procedures can solve these items. Besides that, six of these items are loaded to factor 1 according to exploratory factor analysis results.

Group7: This group is named as **Advanced Calculations II**. Items MT2.12, MT2.13, MT2.14, MT2.15, MT2.16, MT2.17, MT2.18, MT2.19, MT2.20 and MT2.21 have common properties to be in this group. In all of these items, result of an advanced calculation (high school mathematics topics after limit of a function subject) is asked. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out routine algebraic procedures can solve these items. Besides that, all of these items are loaded to factor 1 according to exploratory factor analysis results.

Group8: This group is named as **Advanced Generalizations**. Items MT2.5 and MT2.8 have common properties to be in this group. In these items, some generalization by understanding given situation is asked and there is not any calculation in this decision. In all of these items, content is related to high school mathematics topics between trigonometry and limit of a function subject. Similar to Bloom and TIMSS 2007 cognitive process definitions (Comprehension and Knowing), students who have ability to follow line of reasoning can solve these items. Besides that, MT2.8 is loaded to factor 1 and MT2.5 is loaded to factor 5 according to exploratory factor analysis results.

Group9: This group is named as **Triangle Calculations**. Items MT1.21, MT1.22, MT2.23, MT2.24 and MT2.25 have common properties to be in this group. In all of these items, calculation of an angle, an area or a length related to a triangle is asked. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out algorithms related to Triangles can solve these items. Besides that, four of these items are loaded to factor 2 according to exploratory factor analysis results.

Group10: This group is named as **Quadrangle Calculations**. Items MT1.23, MT2.26 and MT2.27 have common properties to be in this group. In all of these items, a calculation of an angle, an area or a length related to a quadrangle is asked. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out algorithms related to Quadrangles can solve these items. Besides that, all of these items are loaded to factor 2 according to exploratory factor analysis results.

Group11: This group is named as **Circle Calculations**. Items MT1.24, MT1.26, MT2.28, MT2.29 and MT2.30 have common properties to be in this group. In all of these items, a calculation of an angle, an area or a length related to a circle is asked. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out algorithms related to Circles can solve these items. Besides that, four of these items are loaded to factor 2 according to exploratory factor analysis results.

Group12: This group is named as **Analytic Geometry Calculations**. Items MT1.29 and MT1.30 have common properties to be in this group. In all of these items, a calculation related to a coordinate plane is asked. Similar to Bloom and TIMSS 2007 cognitive process definitions (Computation and Knowing), students who have ability to carry out algorithms related to Coordinate Plane can solve these items. Besides that, all of these items are loaded to factor 2 according to exploratory factor analysis results.

Group13: This group is named as **Geometry Problems**. Items MT1.25, MT1.27, MT1.28 and MT2.22 have common properties to be in this group. In all of these items, a short passage which defines situation related geometry is given and by writing a mathematical expression related to this passage, performing steps of calculations is expected. Similar to Bloom and TIMSS 2007 cognitive process definitions (Application and Applying), students who have ability to solve routine problems related to content of Geometry can solve these items. Besides that, all of these items are loaded to factor 2 according to exploratory factor analysis results.

Out of thirteen groups, nine of item groups are related to “Computation/Knowing” cognitive process, two of item groups are related to

“Comprehension/Knowing” cognitive process and two of item groups are related to “Application/Analysis” cognitive process. None of the item group is related to “Analysis/Reasoning” cognitive process.

Table 4.4 Items in each groups-2006 SST

Groups	Items in groups
Basic Calculations	MT1.1, MT1.2, MT1.3, MT1.14
Symbolic Calculations I	MT1.4, MT1.5, MT1.6, MT1.7, MT1.11, MT1.12, MT1.13
Generalizations	MT1.9, MT1.10, MT1.20
Word Problems	MT1.8, MT1.15, MT1.16, MT1.17, MT1.18, MT1.19
Symbolic Calculations II	MT2.1, MT2.2
Advanced Calculations I	MT2.3, MT2.4, MT2.6, MT2.7, MT2.9, MT2.10, MT2.11
Advanced Calculations II	MT2.12, MT2.13, MT2.14, MT2.15, MT2.16, MT2.17, MT2.18, MT2.19, MT2.20, MT2.21
Advanced Generalizations	MT2.5, MT2.8
Triangle Calculations	MT1.21, MT1.22, MT2.23, MT2.24, MT2.25
Quadrangle Calculations	MT1.23, MT2.26, MT2.27
Circle Calculations	MT1.24, MT1.26, MT2.28, MT2.29, MT2.30
Analytic Geometry Calculations	MT1.29, MT1.30
Geometry Problems	MT1.25, MT1.27, MT1.28, MT2.22

4.2.2 Confirmatory Factor Analysis of 2006 SST

Confirmatory factor analysis is performed to answer the second research question of “what are the dimensions of mathematics subtests in the student selection tests” by identifying number of factors in SST mathematics sections. Exploratory factor analysis results indicated that there are five dimensions in SST mathematics sections. By using item groups, a five-factor model will be tested. If fit values of five-factor model are not satisfactory, four-factor, three-factor, two-factor and one-factor models will be tested. LISREL 8.7 is used to test fit of several competitive models and to confirm the proposed relations of observed variables with latent variables.

In order to have valid SEM or CFA results, multivariate normality and linearity assumption should hold. Besides, observed variables should be continuous and interval scaled. Ratio of sample size to number of variable is also important. As it is shown in Appendix I, multivariate normality assumption is almost met; all skewness values and kurtosis values are within acceptable values, except Advanced Calculation II and Advanced Generalizations. George and Mallery (2001) stated that skewness and kurtosis value between ± 1 is considered as excellent and between ± 2 is considered acceptable for normality. Also, in this dataset, there are no missing data, no outliers, and data is continuous. With use of summated scales, scores are continuous. For 2006 data, there are 872 956 participants and 13 observed variables and the ratio of participant to observed variable is 67150:1 which is very huge.

Hypothesis about the number of factors is tested in two stages. First, competing, nested hypothesized models are tested to determine the number of factors in the SST Mathematics sections. In second stage, by eliminating items with low corrected item-total correlations, the accepted model in first stage is retested to investigate whether any improvement occurs.

All of five nested competitive models consist of the following thirteen observed variables: Basic Calculations, Symbolic Calculations I, Generalizations, Word Problems, Symbolic Calculations II, Advanced Calculations I, Advanced Calculations II, Advanced Generalizations, Triangle Calculations, Quadrangle Calculations, Circle Calculations, Analytic Geometry Calculations and Geometry Problems. These five hypothesized models are described below:

Model 1: There are **five correlated dimensions**; *Basic Computation Ability*, made up of three item groups; *Generalization* Ability, made up of two item groups; *Advanced Computation Ability*, made up of two item groups; *Problem Solving Ability*, made up of two item groups, and *Geometry Ability*, made up of four item groups.

Model 2: There are **four correlated dimensions**; *Basic Computation Ability*, made up of four item groups; *Advanced Computation Ability*, made up of three item groups; *Problem Solving Ability*, made up of two item groups, and *Geometry Ability*, made up of four item groups.

Model 3: There are **three correlated dimensions**; *Basic Computation Ability*, made up of five item groups; *Advanced Computation Ability*, made up of three item groups; and *Geometry Ability*, made up of five item groups.

Model 4: There are **two correlated dimensions**; *Computation Ability*, made up of eight item groups, and *Geometry Ability*, made up of five item groups.

Model 5: There is only **one dimension**, made up of thirteen item groups, in which it is claimed that mathematics section of SST measures only a *General Mathematical Ability*.

Schumacker and Lomax (2004) proposed five steps to perform confirmatory factor analyses; model specification, model identification, model estimation, model testing and model modification. For these five different models these steps are followed. For model estimation method, the Maximum Likelihood (ML) estimation technique is used for all models, because, if the observed variables are interval scaled and multivariate normal, then the ML estimates, standard errors, and chi-square test are appropriate. For final step in confirmatory factor analysis, if the model has poor model fit indices, it is necessary to make changes to a specified model (model modification). In this study this step is not used because, the best fitting model will be chosen among several competitive models.

4.2.2.1 Model 1: SST Mathematics sections have five-factor structure

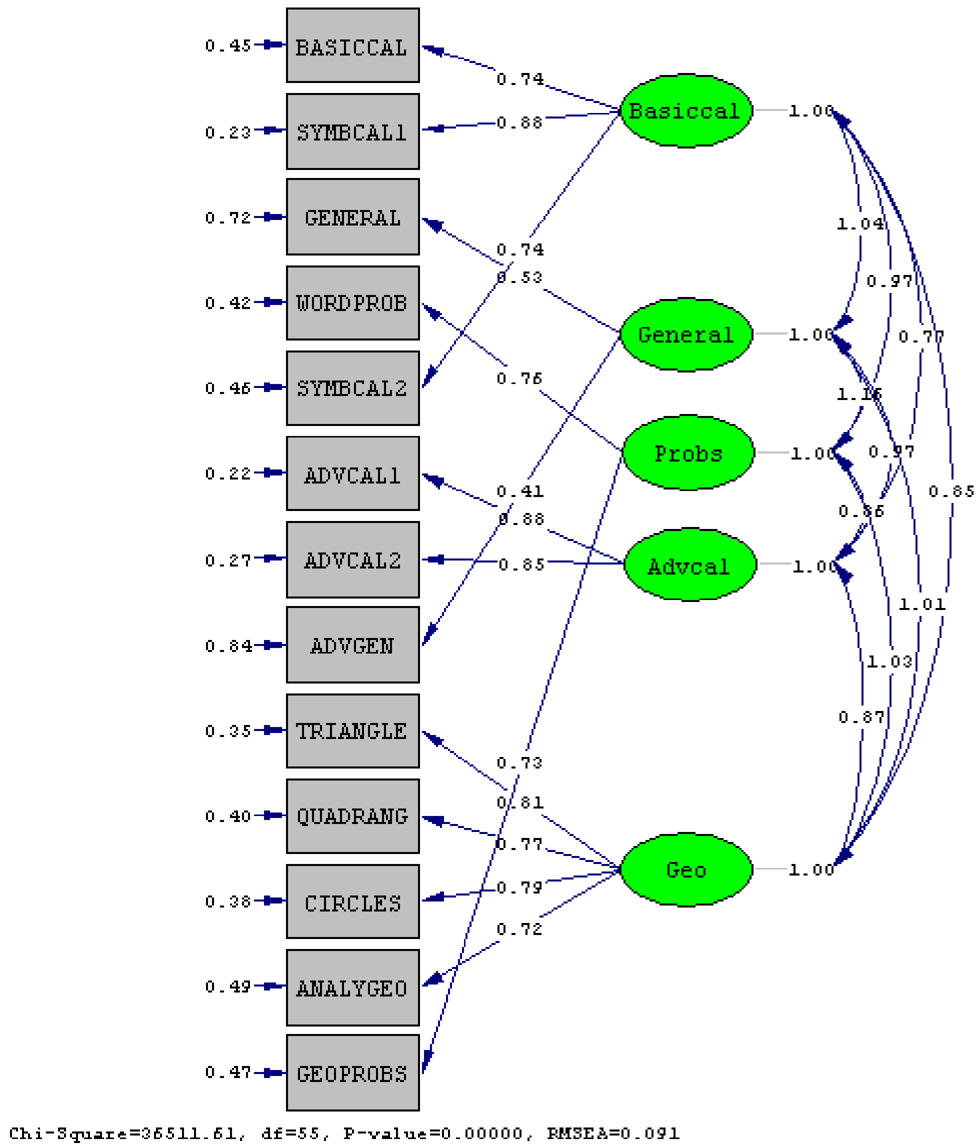
LISREL diagram of the proposed theoretical model is shown in Figure 4.1. The observed variables are shown by rectangles and the latent variables are shown by circle. The measurement errors are shown by arrows to observed variables on the left and show that some portion of each observable variable is measuring something other than the hypothesized factor. A curved, double-headed line between latent variables (for one-factor model there is one latent variable) means that they have shared variance or are correlated with no implied direction of effect. A line with one arrow directed from a factor to an observed variable shows the relation between that factor and that measure. These relationships are interpreted as factor loadings. For clarity in the text, italic letters will be used for names of latent variables (factors) (Schumacker & Lomax, 2004).

Model Specification: Five-factor model

In this five-factor model, Basic Calculations, Symbolic Calculations I, and Symbolic Calculations II are hypothesized to measure *Basic Computation Ability* (Basiccal); Generalization and Advanced Generalizations are hypothesized to measure *Generalization Ability* (General); Advanced Calculations I, and Advanced Calculations II are hypothesized to measure *Advanced Computation Ability* (Advcal); Word Problems and Geometry Problems are hypothesized to measure *Problem Solving Ability* (Probs); and Triangle Calculations, Quadrangle Calculations, Circle

Calculations, Analytic Geometry Calculations are hypothesized to measure *Geometry Ability* (Geo). *Basic Computation Ability*, *Generalization Ability*, *Advanced Computation Ability*, *Problem Solving Ability* and *Geometry Ability* are latent variables that are not directly measured but rather assessed indirectly using proposed observed variables above.

Figure 4.1 Measurement Model for SST Mathematics Sections



Model Identification: Five-factor model

In this model, df is equal to 55. With thirteen variables there are (13(13+1))/2=91 data points and 36 parameters to be estimated (8 regression coefficients, 18 variances, 10 covariances) therefore, according to the order condition, this model is overidentified. Also, according to LISREL, each parameter in this model can be estimated from the covariance matrix therefore, this measurement model is identified.

Model Testing: Five-factor model

The next step is to analyze the fit of confirmatory factor model. If the fit of this model is good, then this model is supported by the sample data. There are many model fit indices which are reported by LISREL and by using these indices which model is most suitable will be decided. A good fitting model has consistent fit indices generally (Ullman, 2001). In these study χ^2, χ^2/df, GFI, AGFI, RMR, RMSEA, NFI, RFI, IFI, CFI, and ECVI values will be reported for each separate model and these values will be compared.

The chi-square, χ^2= 36511.431, is significant with df=55, and the significance level is p=0.00. However, the χ^2 value is highly affected from sample size; as the sample size increases so does χ^2 (Schumacker & Lomax, 2004). Another criteria directly related with chi-square is the ratio of the chi-square value to the degrees of freedom which is 663.844 in this model. Although this value is very large than expected value of 2, it is impossible to have nonsignificant chi-square related values with sample size of 872 956.

The independence chi-square test value is 1699744.237 with degrees of freedom of 78 and this test is significant. Significant independence chi-square test means that there is some relationship among variables as desired.

The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI) of this model is 0.934 and 0.891 respectively. Since GFI value is lower than 0.95 and AGFI is lower than 0.90, this model has poor fit to the data.

The Root-Mean-Square Residual (RMR) of this model is 0.092 which is higher than 0.05 and close to 0.10, therefore indicating moderate fit to the data. Also, the Root-Mean-Squared Error of Approximation (RMSEA) of the model is 0.091 which is higher than 0.06 and close to 0.10, therefore shows a sign of moderate fit to the data. Also, the 90 percent confidence interval for RMSEA is 0.090 and 0.092.

The Normed Fit Index (NFI), The Relative Fit Index (RFI), The Incremental Fit Index (IFI) and the Comparative Fit Index (CFI) of this model are 0.980, 0.972, 0.980 and 0.980 respectively. Although these values are higher than criteria value of 0.95, some of the previous fit indices were showing poor or moderate fit. Therefore, NFI, RFI, IFI and CFI values of this model will be compared with NFI, RFI, IFI and CFI values of other models.

The expected cross-validation index (ECVI) is 0.457, which is very useful in the comparing different models and deciding how many factors is measured by SST mathematics sections. A model with the smallest ECVI value shows the greatest possibility to cross-validate. Therefore, all ECVI values of competitive models will be compared at the end of section.

Confirmatory factor analysis for five-factor model is performed to answer the second research question of “what are the dimensions of mathematics subtests in the student selection tests”. As the fit indices for five-factor model generally indicate poor or moderate fit to the data, also, other models will be tested by reducing number of factor according to low factor loadings or high correlations among dimensions. In following section, four-factor model will be tested by reducing one dimension from five-factor model. Process of reducing number of factors and testing new models will continue in order to decide the dimensions of mathematics subtests. It is aimed to find out best fitting model to the data.

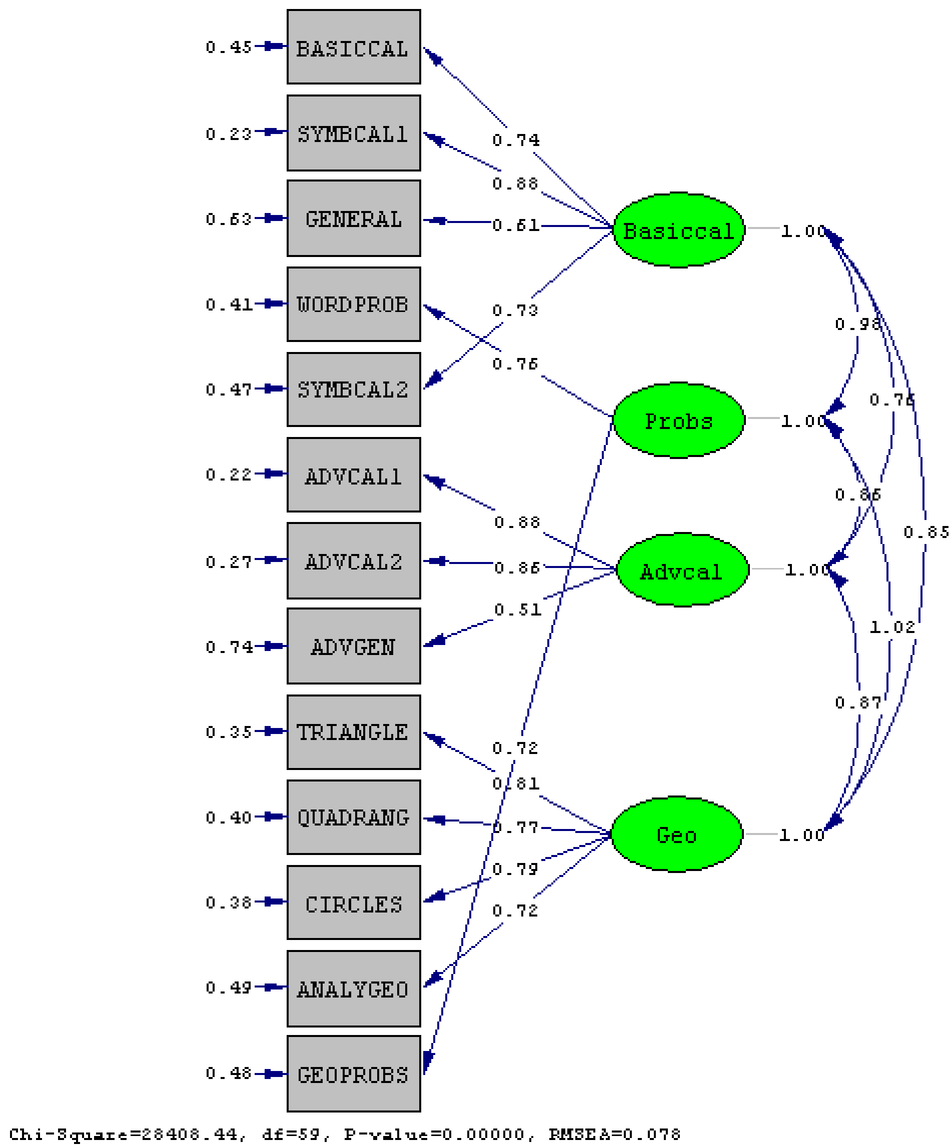
4.2.2.2 Model 2: SST Mathematics sections have four-factor structure

As five-factor model generally have poor or moderate fit, a new model is formed. This new four-factor model is composed by eliminating *Generalization Ability* from five-factor model. This elimination is done on the basis of low factor loadings of Generalization-*Generalization Ability* and Advanced Generalization-*Generalization Ability* (0.53 and 0.41). In four-factor model Generalization is hypothesized to be related to *Basic Computation Ability* and Advanced Generalization is hypothesized to be related to *Advanced Computation Ability*.

Model Specification: Four-factor model

In this four-factor model, Basic Calculations, Symbolic Calculations I, Generalizations and Symbolic Calculations II are hypothesized to measure *Basic Computation Ability* (Basiccal); Advanced Calculations I, Advanced Calculations II and Advanced Generalizations are hypothesized to measure *Advanced Computation Ability* (Advcal); Word Problems and Geometry Problems are hypothesized to measure *Problem Solving Ability* (Probs); and Triangle Calculations, Quadrangle Calculations, Circle Calculations, Analytic Geometry Calculations are hypothesized to measure *Geometry Ability* (Geo). *Basic Computation Ability*, *Advanced Computation Ability*, *Problem Solving Ability* and *Geometry Ability* are latent variables that are not directly measured but rather assessed indirectly using proposed observed variables above. The LISREL diagram of the proposed theoretical model is shown in Figure 4.2.

Figure 4.2 Measurement Model for SST Mathematics Sections



Model Identification: Four-factor model

In this model, df is equal to 59. With thirteen variables there are (13(13+1))/2=91 data points and 32 parameters to be estimated (9 regression coefficients, 17 variances, 6 covariances) therefore, according to the order condition, this model is overidentified. Also, according to LISREL, each parameter in this model can be estimated from the covariance matrix therefore, this measurement model is identified.

Model Testing: Four-factor model

The chi-square, χ^2 = 28408.437, is significant with df=59, and the significance level is p=0.00. However, the χ^2 value is highly affected from sample size; as the sample size increases so does χ^2 (Schumacker & Lomax, 2004). Another criteria directly related with chi-square is the ratio of the chi-square value to the degrees of freedom which is 481.498 in this model. Although this value is very large than expected value of 2, it is impossible to have nonsignificant chi-square related values with sample size of 872 956.

The independence chi-square test value is 1699744.237 with degrees of freedom of 78 and this test is significant, therefore, significant independence chi-square test means that there is some relationship among variables as desired.

The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI) of this model is 0.948 and 0.920 respectively. Since GFI value is close to 0.95 and AGFI is higher than 0.90, it is secure to say that this model has a good fit to the data for this fit indices.

The Root-Mean-Square Residual (RMR) of this model is 0.086 which is higher than 0.05 and less than 0.10, therefore indicating a moderate fit to the data.

Also, the Root-Mean-Squared Error of Approximation (RMSEA) of the model is 0.078 which is higher than 0.06 and less than 0.10, therefore shows a sign of moderate fit to the data. Also, the 90 percent confidence interval for RMSEA is 0.077 and 0.078.

The Normed Fit Index (NFI), The Relative Fit Index (RFI), The Incremental Fit Index (IFI) and the Comparative Fit Index (CFI) of this model are 0.984, 0.978, 0.984 and 0.984 respectively. Although these values are higher than criteria value of 0.95, some of the previous fit indices were showing moderate fit. Therefore, NFI, RFI, IFI and CFI values of this model will be compared with NFI, RFI, IFI and CFI values of other models.

The expected cross-validation index (ECVI) is 0.356, which is very useful in the comparing different models and deciding how many factors is measured by SST mathematics sections. A model with the smallest ECVI value shows the greatest possibility to cross-validate. Therefore, all ECVI values of competitive models will be compared at the end of section.

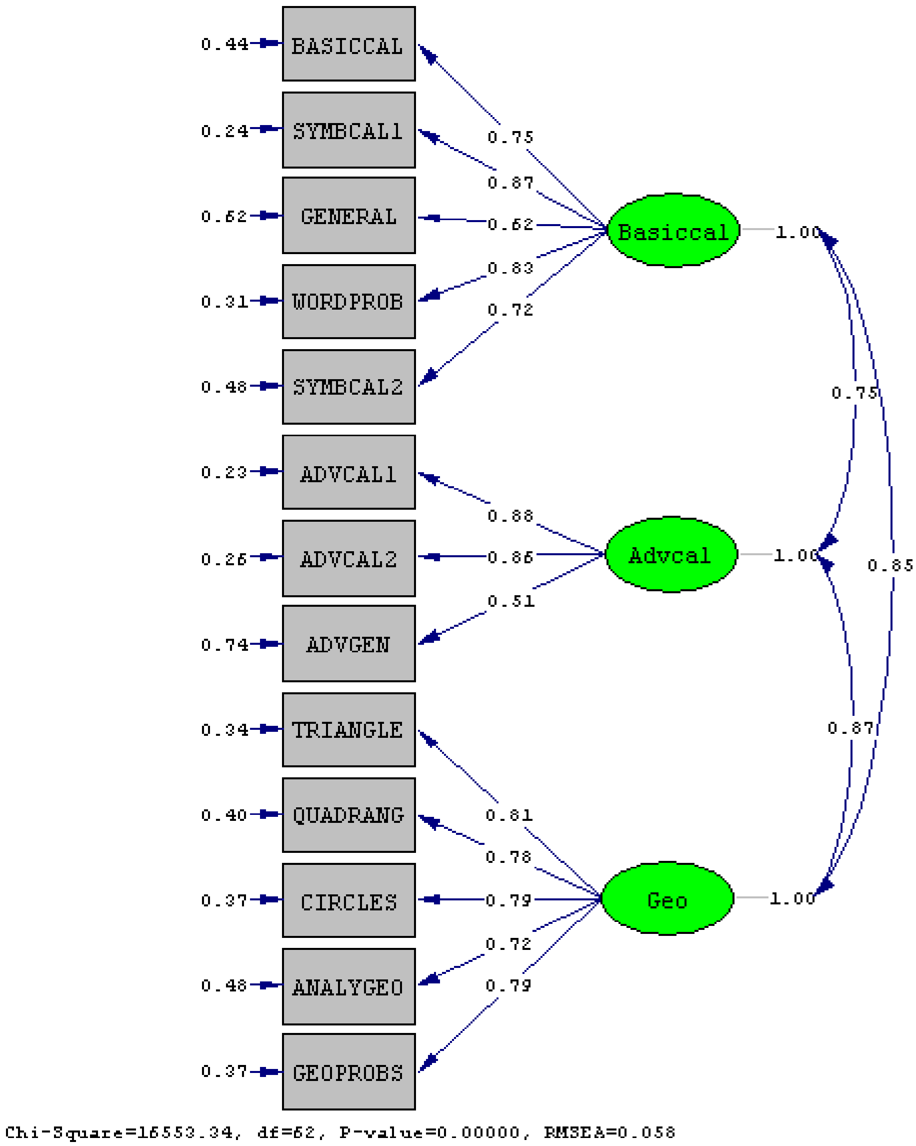
4.2.2.3 Model 3: SST Mathematics sections have three-factor structure

As four-factor model have moderate fit, a new model is formed. This new three-factor model is composed by eliminating *Problem Solving Ability* from four-factor model. This elimination is done on the basis of very high correlation between *Problem Solving Ability* and *Basic Computation Ability* (0.98). In three-factor model Word Problem is hypothesized to be related to *Basic Computation Ability* and Geometry Problems is hypothesized to be related to *Geometry Ability*.

Model Specification: Three-factor model

In this three-factor model, Basic Calculations, Symbolic Calculations I, Generalizations, Word Problems, and Symbolic Calculations II are hypothesized to measure *Basic Computation Ability* (Basiccal); Advanced Calculations I, Advanced Calculations II and Advanced Generalizations are hypothesized to measure *Advanced Computation Ability* (Advcal); and Triangle Calculations, Quadrangle Calculations, Circle Calculations, Analytic Geometry Calculations and Geometry Problems are hypothesized to measure *Geometry Ability* (Geo). *Basic Computation Ability*, *Advanced Computation Ability* and Geo are latent variables that are not directly measured but rather assessed indirectly using proposed observed variables above. The LISREL diagram of the proposed theoretical model is shown in Figure 4.3.

Figure 4.3 Three-factor Measurement Model for SST Mathematics Sections



Model Identification: Three-factor model

In this model, df of this model is equal to 62. With 13 variables there are (13(13+1))/2=91 data points and 29 parameters to be estimated (10 regression coefficients, 16 variances and 3 covariances), therefore, according to the order condition, this model is overidentified. Also, according to LISREL, each parameter in this model can be estimated from the covariance matrix; therefore, this measurement model is identified.

Model Testing: Three-factor model

The chi-square, χ^2 = 16553.342, is significant with df=62, and the significance level is p=0.00. However, the χ^2 value is highly affected from sample size; as the sample size increases so does χ^2 (Schumacker & Lomax, 2004). Another criteria directly related with chi-square is the ratio of the chi-square value to the degrees of freedom which is 266.989 in this model. Although this value is very large than expected value of 2, it is impossible to have nonsignificant chi-square related values with sample size of 872 956.

The independence chi-square test value is 1699744.237 with degrees of freedom of 78 and this test is significant. Significant independence chi-square test means that there is some relationship among variables as desired.

The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI) of this model is 0.969 and 0.955 respectively. Since GFI value is higher than 0.95 and AGFI is higher than 0.90, it is secure to say that three-factor model has a good fit to the data.

The Root-Mean-Square Residual (RMR) of this model is 0.054 which is close to 0.05, therefore indicating a good fit to the data. Also, the Root-Mean-Squared

Error of Approximation (RMSEA) of the model is 0.058, lower than 0.06, shows a good fit to the data. Also, 90 percent confidence interval for RMSEA is 0.057 and 0.058.

The Normed Fit Index (NFI), The Relative Fit Index (RFI), The Incremental Fit Index (IFI) and the Comparative Fit Index (CFI) of this model are 0.991, 0.988, 0.991, and 0.991 respectively. All of these indices are indication of a good fitting model. Besides that NFI, RFI, IFI and CFI values of this model will be compared with NFI, RFI, IFI and CFI values of other models.

The expected cross-validation index (ECVI) is 0.208, which is very useful in the comparing different models and deciding how many factors is measured by SST mathematics sections. A model with the smallest ECVI value shows the greatest possibility to cross-validate. Therefore, all ECVI values of competitive models will be compared at the end of section.

4.2.2.4 Model 4: SST Mathematics sections have two-factor structure

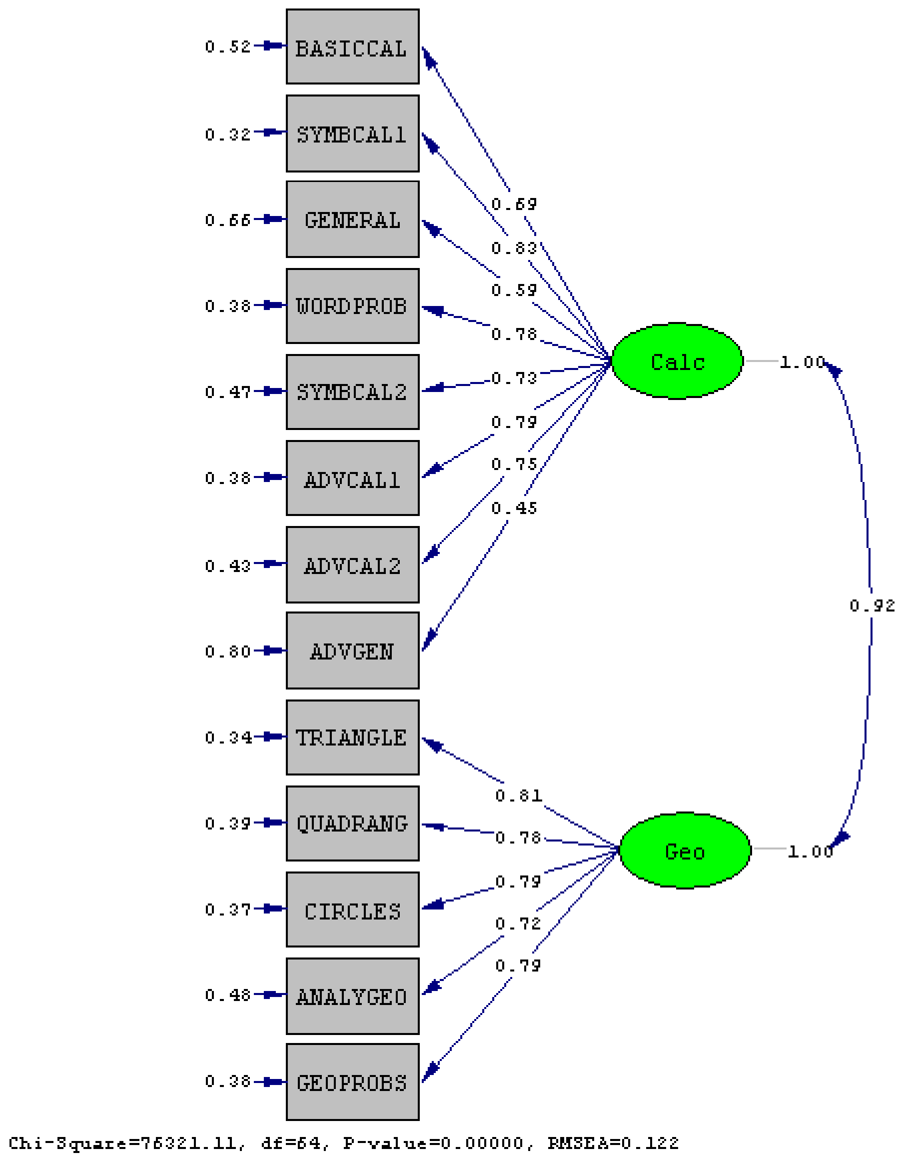
Three-factor model have good fit to the data. However, in order to get better model, two-factor model also formed. This new two-factor model is composed by combining *Basic Computation Ability* and *Advanced Computation Ability* from five-factor model.

Model Specification: Two-factor model

In this two-factor model, Basic Calculations, Symbolic Calculations I, Generalizations, Word Problems, Symbolic Calculations II, Advanced Calculations I, Advanced Calculations II and Advanced Generalizations are hypothesized to measure *Computation Ability* (Calc); whereas, Triangle Calculations, Quadrangle Calculations, Circle Calculations, Analytic Geometry Calculations and Geometry

Problems are hypothesized to measure *Geometry Ability* (Geo). *Computation Ability* and *Geometry Ability* are latent variables that are not directly measured but rather assessed indirectly using proposed observed variables above. The LISREL diagram of the proposed theoretical model is shown in Figure 4.4.

Figure 4.4: Two-factor Measurement Model for SST Mathematics Sections



Model Identification: Two-factor model

In this model, df of this model is equal to 64. With thirteen variables there are (13(13+1))/2=91 data points and 27 parameters to be estimated (11 regression coefficients, 15 variances and 1 covariances), therefore, according to the order condition, this model is “overidentified”. Also, according to LISREL, each parameter in this model can be estimated from the covariance matrix; therefore, this measurement model is identified.

Model Testing: Two-factor model

The chi-square, χ^2 = 76321.115, is significant with df=64, and the significance level is p=0.00. However, the χ^2 value is highly affected from sample size; as the sample size increases so does χ^2 (Schumacker & Lomax, 2004). Another criteria directly related with chi-square is the ratio of the chi-square value to the degrees of freedom which is 1192.517 in this model. Although this value is very large than expected value of 2, it is impossible to have nonsignificant chi-square related values with sample size of 872 956.

The independence chi-square test value is 1699744.237 with degrees of freedom of 78 and this test is significant. Significant independence chi-square test means that there is some relationship among variables as desired.

The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI) of this model is 0.872 and 0.818 respectively. Since GFI value is lower than 0.95 and AGFI is lower than 0.90, it is secure to say that two-factor model have a poor fit to the data.

The Root-Mean-Square Residual (RMR) of this model is 0.136 which is higher than 0.05, therefore indicating a poor fit to the data. Also, the Root-Mean-

Squared Error of Approximation (RMSEA) of the model is 0.122 shows a poor fit to the data. Also, 90 percent confidence interval for RMSEA is 0.121 and 0.123.

The Normed Fit Index (NFI), The Relative Fit Index (RFI), The Incremental Fit Index (IFI) and the Comparative Fit Index (CFI) of this model are 0.967, 0.960, 0.967, and 0.967 respectively. Although these values are higher than criteria value of 0.95, previous fit indices were showing poor fit. Therefore, NFI, RFI, IFI and CFI values of this model will be compared with NFI, RFI, IFI and CFI values of other models.

The expected cross-validation index (ECVI) is 0.955, which is very useful in the comparing different models and deciding how many factor is measured by SST mathematics sections. A model with the smallest ECVI value shows the greatest possibility to cross-validate. Therefore, all ECVI values of competitive models will be compared at the end of section.

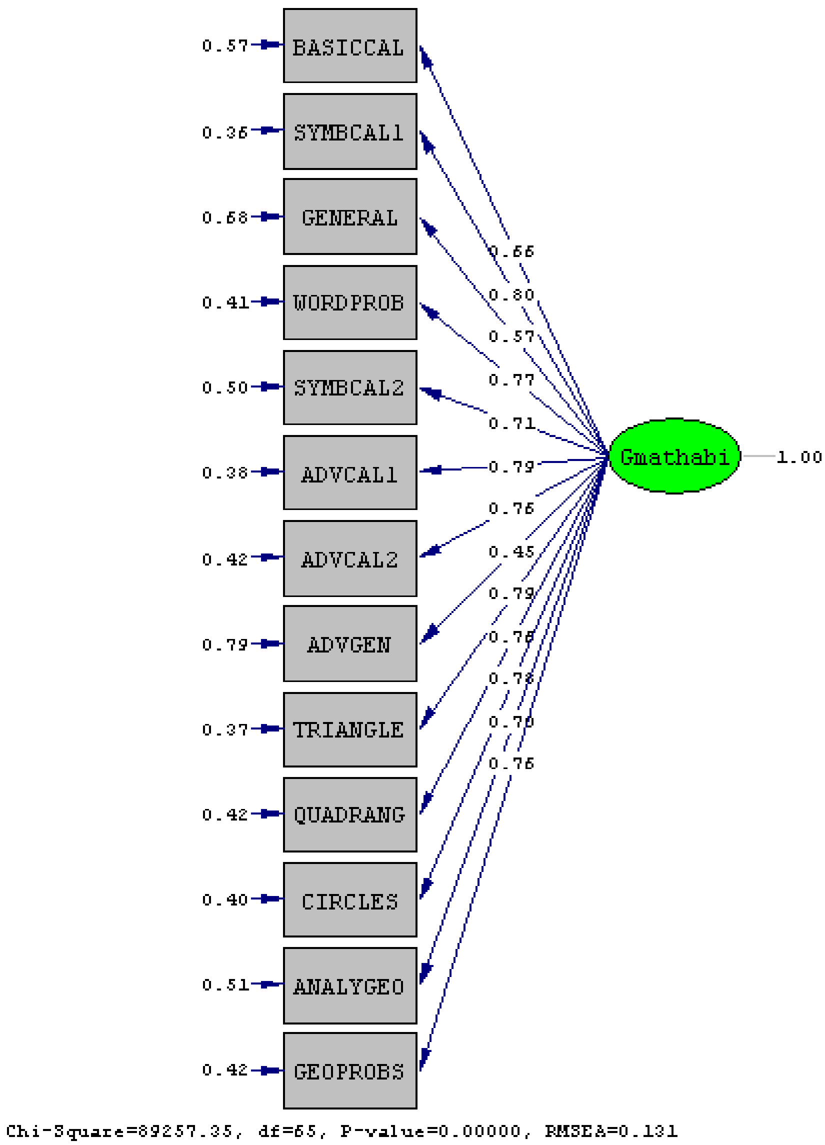
4.2.2.5 Model 5: SST Mathematics sections have one-factor structure

Two-factor model have poor fit to the data. As a final model one-factor model is formed and tested.

Model Specification: One-factor model

In this one-factor model, Basic Calculations, Symbolic Calculations I, Generalizations, Word Problems, Symbolic Calculations II, Advanced Calculations I, Advanced Calculations II, Advanced Generalizations, Triangle Calculations, Quadrangle Calculations, Circle Calculations, Analytic Geometry Calculations and Geometry Problems are hypothesized to measure *General Mathematical Ability* (Gmathabi). *General Mathematical Ability* is a latent variable that is not directly measured but rather assessed indirectly using proposed observed variables above. The LISREL diagram of the proposed theoretical model is shown in Figure 4.5.

Figure 4.5 One-factor Measurement Model for SST Mathematics Sections



Model Identification: One-factor model

In this model, df is equal to 65. With thirteen variables there are (13(13+1))/2=91 data points and 26 parameters will be estimated (12 regression coefficients, 14 variances). Therefore, according to the order condition, this model is "overidentified". The other condition requires algebraically determine whether each parameter in the model can be estimated uniquely from the covariance matrix. According to LISREL, each parameter in this model can be estimated from the covariance matrix; therefore, this measurement model is identified.

Model Testing: One-factor model

The chi-square, $\chi^2 = 89257.350$, is significant with df=65, and the significance level is p=0.00. However, the χ^2 value is highly affected from sample size; as the sample size increases so does χ^2 (Schumacker & Lomax, 2004). Another criteria directly related with chi-square is the ratio of the chi-square value to the degrees of freedom which is 1373.19 in this model. Although this value is very large than expected value of 2, it is impossible to have nonsignificant chi-square related values with sample size of 872 956.

The independence chi-square test value is 1699744.237 with 78 degrees of freedom and this test is significant. Significant independence chi-square test means that there is some relationship among variables as desired.

The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI) of this model is 0.853 and 0.795 respectively. Since GFI value is lower than 0.95 and AGFI is lower than 0.90, it is secure to say that one-factor model have a poor fit to the data.

The Root-Mean-Square Residual (RMR) of this model is 0.143 which is higher than 0.05, therefore indicating a poor fit to the data. Also, the Root-Mean-Squared Error of Approximation (RMSEA) of the model is 0.131 and shows a poor fit to the data. Also, 90 percent confidence interval for RMSEA is 0.130 and 0.132.

The Normed Fit Index (NFI), the Relative Fit Index (RFI), the Incremental Fit Index (IFI) and the Comparative Fit Index (CFI) of this model are 0.961, 0.953, 0.961, and 0.961 respectively. Although these values are higher than criteria value of 0.95, previous fit indices were showing poor fit. Therefore, NFI, RFI, IFI and CFI values of this model will be compared with NFI, RFI, IFI and CFI values of other models.

The expected cross-validation index (ECVI) is 1.116, which is very useful in the comparing different models and deciding how many factor is measured by SST mathematics sections. A model with the smallest ECVI value shows the greatest possibility to cross-validate.

4.2.2.6 Comparison of hypothesized models

Table 4.5 shows values of selected eleven fit indices for proposed five different models. For these comparisons, higher values of GFI, AGFI, NFI, RFI, IFI, CFI and lower values of χ^2, χ^2/df, RMR, RMSEA, ECVI are sign of better fit. As it is explained in Chapter 3, GFI, NFI, RFI, IFI, CFI values larger than 0.95; AGFI value larger than 0.90; RMR value less than 0.05; RMSEA value less than 0.06 shows good fit. RMR and RMSEA values higher than 0.10 means poor fit. These criteria will be used to evaluate fit of competitive models. By comparing values of that many indices, best model will be chosen more confidently.

Fit indices of five-factor model do not supported by data. Four-factor model and three-factor model seem plausible with having generally good fit indices values. However, three-factor model have better values than four-factor model. Three-factor model have lower $\chi^2, \chi^2 / df$, RMR, RMSEA and ECVI values. In all of these five types of value, lower values mean better fit. Besides that, three-factor model have higher GFI, AGFI, NFI, RFI, IFI, CFI values. In all of these six types of value, higher values mean better fit. It is important to note that, there is not any fit index in which four-factor model has better fitting value. Therefore, three-factor model is selected as a good representation of what SST Mathematical section actually measures and this model is expected to cross-validate with 2007 SST mathematics sections data also. Two-factor model and one-factor model are not supported by data also.

The second research question asks "What are the dimensions of mathematics subtests (first and second section) in the student selection tests?". As three-factor model is supported by the data, there are three dimensions of mathematics subtests in SST: one of them measures *Basic Computation Ability*; one of them measures *Advanced Computation Ability*; and the other measures *Geometry Ability*. This result shows that SST is good in measuring "Computation/Knowing" cognitive process in which *Basic Computation Ability*, *Advanced Computation Ability* and *Geometry Ability are* mainly related to reaching successfully to a result by performing some steps of calculation.

On the other hand, in five-factor model it was proposed that *Generalization Ability* (related to "Comprehension/Knowing" cognitive process) and *Problem Solving Ability* (related to "Application/Analysis" cognitive process) are also measured in addition to three-factor model. Similarly, in four-factor model it was proposed that *Problem Solving Ability* (related to "Application/Analysis" cognitive process) is also measured in addition to three-factor model. However, these two models are not accepted because of poorer fit indices. Therefore, it is concluded that generalization and problem solving items are not functioning as it is aimed. These items measures some sort of computation ability. Consequently, these results showed that SST mathematics section do not measure beyond computation abilities.

Table 4.5 Comparison of Fit Indices for hypothesized models-2006 SST

	Five-factor Model	Four-factor Model	Three-factor Model	Two-factor Model	One-factor Model
x^2	36511.431	28408.437	16553.342	76321.115	89257.350
x^2/df	663. 844	481.498	266.989	1192.517	1373.190
GFI	0.934	0.948	0.969	0.872	0.853
AGFI	0.891	0.920	0.955	0.818	0.795
RMR	0.092	0.086	0.054	0.136	0.143
RMSEA	0.091	0.078	0.058	0.122	0.131
NFI	0.980	0.984	0.991	0.967	0.961
RFI	0.972	0.978	0.988	0.960	0.953
IFI	0.980	0.984	0.991	0.967	0.961
CFI	0.980	0.984	0.991	0.967	0.961
ECVI	0.457	0.356	0.208	0.955	1.116

4.2.2.7 Improving Three-factor Model

Related to the second research question, it is decided that 2006 SST mathematics section has better fit to three-factor structure. In Appendix H, it is seen that some items have low corrected total-item correlations (lower than 0.200). It is worth to investigate whether elimination of these items results in better model fit. In this part, three-factor model without poor items and three-factor model with all items will be compared and better model will be decided.

Appendix H shows that items that have reliability values under 0.200 are MT2.5, MT2.13, MT2.18, and MT2.20. Therefore, one item from Advanced Generalizations group (MT2.5) and three items from Advanced Calculations II group (MT2.13, MT2.18, and MT2.20) are eliminated. Total of 56 items will be used for three-factor model without poor items and total of 60 items will be used for three-factor model with all items.

As it is explained before, only items from Advanced Generalizations and Advanced Calculations II are eliminated. Therefore, only factor loading values of these two observed variables have changed. Advanced Generalizations factor loading of 0.51 increased to 0.61 and Advanced Calculations II factor loading of 0.86 increased to 0.87. Therefore elimination of poor items improves factor loading values.

Table 4.6 gives fit indices for these models. Three-factor model with all items and three-factor model without poor items both have good fit indices values. However, three-factor model without poor items have slightly better values than three-factor model with all items. Three-factor model without poor items have lower χ^2, χ^2/df, RMR and ECVI values. In all of these four types of value, lower values mean better fit. Besides that, three-factor model without poor items have higher RFI value. In this index, higher values mean better fit. It is important to note that, there is not any fit index in which three-factor model with all items has better fitting value than three-factor model without poor items. Therefore, elimination of poor items affects positively the fit indices values. Without these items, there is a better fit to the model.

Table 4.6 Comparison of Fit Indices for Three-factor model with all items and Three-factor model without poor items -2006 SST

	Three-factor Model with all items	Three-factor Model without poor items
x^2	16553.342	16517.707
x^2/df	266.989	266.414
GFI	0.969	0.969
AGFI	0.955	0.955
RMR	0.054	0.050
RMSEA	0.058	0.058
NFI	0.991	0.991
RFI	0.988	0.989
IFI	0.991	0.991
CFI	0.991	0.991
ECVI	0.208	0.207

4.2.2.8 Standardized Path Coefficients

As final step for the second research question standardized path coefficient will be analyzed. If the established model fits the data, it is necessary to examine the statistically significant relationships within the model. It is confirmed that three-factor structure model without poor items is better model. From now on, all analysis will be conducted with using this model.

Table 4.7 shows standardized estimates (loadings) of each observed variables on related latent variables. Each of the relation between observed variables and latent variables are significant ($p<0.05$) and values are higher than 0.50 which indicates good relations (Stricker et al., 2005). Therefore, it is concluded that Basic Calculations, Symbolic Calculations I, Generalizations, Word Problems and Symbolic Calculations II is a significant indicator of *Basic Computation Ability*; Advanced Calculations I, Advanced Calculations II and Advanced Generalizations is a significant indicator of *Advanced Computation Ability*; Triangle Calculations, Quadrangle Calculations, Circle Calculations, Analytic Geometry Calculations and Geometry Problems is a significant indicator of *Geometry Ability*. Therefore, it can be stated more confidently that SST have three-dimensions.

Table 4.7 Standardized Path Coefficients-2006 SST

Observed Variables	Latent Variables	Standardized Estimates
Basic Calculations	*Basic Computation Ability*	0.75
Symbolic Calculations I	*Basic Computation Ability*	0.87
Generalizations	*Basic Computation Ability*	0.62
Word Problems	*Basic Computation Ability*	0.83
Symbolic Calculations II	*Basic Computation Ability*	0.72
Advanced Calculations I	*Advanced Computation Ability*	0.88
Advanced Calculations II	*Advanced Computation Ability*	0.87
Advanced Generalizations	*Advanced Computation Ability*	0.61
Triangle Calculations	*Geometry Ability*	0.81
Quadrangle Calculations	*Geometry Ability*	0.78
Circle Calculations	*Geometry Ability*	0.79
Analytic Geometry Calculations	*Geometry Ability*	0.72
Geometry Problems	*Geometry Ability*	0.79

4.3 Cross Validation of Results across Gender Groups and School Types

The third research question asks “Do dimensions of 2006 SST mathematics subtests provide stable factorial structure across gender groups and school types?”. In order to answer this question, multiple group analysis related to accepted three-factor model for gender groups and school types is conducted.

The purpose of this section is to assess the invariance of the defined and confirmed dimensions in SST Mathematics sections for each of the subgroups defined as a) gender b) school type. The analysis for this section is base on testing whether number of factors, factor correlations, error variances, and factor loadings are invariant across groups.

In order to test invariance of factor structures the following models will be tested for each subgroup.

Model A: The number of factors is invariant,

Model B: The factor correlations are invariant,

Model C: The factor correlations and error variances are invariant,

Model D: The factor correlation, error variances and factor loadings are invariant.

In order to compare models described above, fit indices of df, χ^2, χ^2/df CFI, NFI and RMSEA is reported for overall model and fit indices of SRMR and GFI is reported for each subgroup separately (Stricker et al., 2005; Stricker & Rock, 2008).

4.3.1 Testing Equality of Factor structures across Gender groups

The equality of factor structures for males and females is tested in this section. Table 4.8 shows the fit indices values separately for the Model A, Model B, Model C and Model D which are defined previously. With related to the invariance of the number of factors across males and females, the goodness of fit indexes for the individual samples and for the overall analysis is quite satisfactory. Therefore, the number of factors is invariant for males and females.

With related to the invariance of the factor correlations across males and females, the goodness of fit indexes for the individual samples and for the overall analysis is good. Therefore, the factor correlations are invariant for males and females.

With related to the invariance of the factor correlations and error variances across males and females, the goodness of fit indexes for the individual samples and for the overall analysis is good. Therefore, the factor correlations and error variances are invariant for males and females.

With related to the invariance of the factor correlations, error variances and factor loadings across males and females, the goodness of fit indexes for the individual samples and for the overall analysis is good. Although only SRMR value in this model is higher than other models, it is lower than cut of score of 0.08. Therefore, the factor correlations, error variances and factor loadings are invariant for males and females. Therefore, there is no evidence to reject this model.

Test of equality of factor structure results indicate that SST mathematics sections function similarly for gender groups. This cross validation result is another validity evidence for three-factor structure of SST mathematics sections.

Table 4.8 Tests of Invariance of Factors for Males and Females

	df	x^2	x^2/df	SRMR	GFI	CFI	NFI	RMSEA
A. Number of factors invariant								
Males				0.029	0.970			
Females				0.028	0.971			
Overall	124	31880.148	257.097			0.991	0.991	0.057
B. Factor correlations invariant								
Males				0.031	0.968			
Females				0.030	0.971			
Overall	127	32643.877	257.038			0.991	0.991	0.057
C. Factor correlations and error variances invariant								
Males				0.034	0.966			
Females				0.032	0.968			
Overall	140	35192.247	251.373			0.990	0.990	0.056
D. Factor correlations, error variances and factor loadings invariant								
Males				0.064	0.963			
Females				0.065	0.964			
Overall	153	39380.585	257.389			0.989	0.989	0.057

4.3.2 Testing Equality of Factor structures across School Types

The equality of factor structures for private and public schools is tested in this section. Table 4.9 shows the fit indices values separately for the Model A, Model B, Model C and Model D which are defined previously. With related to the invariance of the number of factors across private and public schools, the goodness of fit indexes for the individual samples and for the overall analysis is good. Therefore, the number of factors is invariant for private and public school students.

With related to the invariance of the factor correlations across private and public schools, the goodness of fit indexes for the individual samples and for the overall analysis is good. Therefore, the factor correlations are invariant for private and public school students.

With related to the invariance of the factor correlations and error variances across private and public schools, the goodness of fit indexes for the individual samples and for the overall analysis is good. Therefore, the factor correlations and error variances are invariant for private and public school students.

With related to the invariance of the factor correlations, error variances and factor loadings across private and public schools, the goodness of fit indexes for the individual samples and for the overall analysis is good. Although only SRMR value in this model is higher than other models, it is almost lower than cut of score of 0.08. Therefore, the factor correlations, error variances and factor loadings are invariant for private and public schools with little unimportant difference in factor loadings. Therefore, there is no evidence to reject this model.

Test of equality of factor structure results indicate that SST mathematics sections function similarly for school types. This cross validation result is another validity evidence for three-factor structure of SST mathematics sections. For multiple group analysis, invariant factor structure across gender groups and school types indicates that SST mathematics section is operating similarly for these groups.

Table 4.9 Tests of Invariance of Factors for Private and Public Schools

	df	x^2	x^2/df	SRMR	GFI	CFI	NFI	RMSEA
A. Number of factors invariant								
Private				0.028	0.970			
Public				0.029	0.969			
Overall	124	32747.906	264.096			0.991	0.991	0.057
B. Factor correlations invariant								
Private				0.028	0.969			
Public				0.029	0.969			
Overall	127	33003.462	259.869			0.990	0.990	0.057
C. Factor correlations and error variances invariant								
Private				0.035	0.961			
Public				0.037	0.957			
Overall	140	44342.569	316.732			0.987	0.987	0.063
D. Factor correlations, error variances and factor loadings invariant								
Private				0.082	0.950			
Public				0.077	0.945			
Overall	153	57317.777	374.625			0.983	0.983	0.068

4.4 Cross Validation of Results with 2007 SST

Fourth research question asks "Do dimensions of 2006 SST mathematics subtests provide stable factorial structure across years?". In order to answer this question, confirmatory factor analysis related to three-factor model for 2007 SST is reconducted. It was confirmed that 2006 SST mathematics sections have three-factor structure in which *Basic Computation Ability*, *Advanced Computation Ability* and *Geometry Ability* is measured. It is important to investigate whether similar results will be obtained with the 2007 SST mathematics data. Same fit indices are used in 2007 data analysis in order to compare both results.

2007 SST mathematics sections also have 60 mathematics items which are administered under two different sections. These items are grouped with using similar process as 2006 SST item grouping. Total of thirteen groups is established. Items in each group are given in Table 4.10.

Table 4.10 Items in each Groups-2007 SST

Groups	Items in groups
Basic Calculations	MT1.1, MT1.2, MT1.3, MT1.4, MT1.6, MT1.7
Symbolic Calculations I	MT1.5, MT1.8, MT1.9, MT1.12, MT1.13, MT1.14, MT1.15
Generalizations	MT1.10, MT1.11, MT1.23
Word Problems	MT1.16, MT1.17, MT1.18, MT1.19, MT1.20, MT1.21, MT1.22
Symbolic Calculations II	MT2.1, MT2.2
Advanced Calculations I	MT2.3, MT2.4, MT2.5, MT2.6, MT2.7, MT2.8, MT2.9, MT2.10, MT2.11
Advanced Calculations II	MT2.12, MT2.14, MT2.15, MT2.16, MT2.18, MT2.19, MT2.20, MT2.21
Advanced Generalizations	MT2.13
Triangle Calculations	MT1.24
Quadrangle Calculations	MT1.25, MT2.22, MT2.22
Circle Calculations	MT1.26, MT1.27, MT2.25, MT2.26, MT2.27, MT2.28
Analytic Geometry Calculations	MT1.29, MT1.30
Geometry Problems	MT1.28, MT2.17, MT2.23, MT2.29, MT2.30

Confirmatory factor analysis results of three-factor model for 2007 SST showed that the chi-square, χ^2 = 32960.696, is significant with df=62, and the significance level is p=0.00. However, the χ^2 value is highly affected from sample size; as the sample size increases so does χ^2 (Schumacker & Lomax, 2004). Another criteria directly related with chi-square is the ratio of the chi-square value to the degrees of freedom which is 531.624 in 2007 SST. Although this value is very large than expected value of 2, it is impossible to have nonsignificant chi-square related values with sample size of 915 161.

The independence chi-square test value is 1699744.237 with degrees of freedom of 78 and this test is significant. Significant independence chi-square test means that there is some relationship among variables as desired.

The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI) of this model is 0.940 and 0.913 respectively. Since GFI value is very close to 0.95 and AGFI is higher than 0.90, it is secure to say that three-factor model has a good fit to the data.

The Root-Mean-Square Residual (RMR) of this model is 0.095 which is higher than 0.05 and close to 0.10, therefore indicating moderate fit to the data. The Root-Mean-Squared Error of Approximation (RMSEA) of the model is 0.081, higher than 0.06 and less than 0.10 shows a moderate fit to the data. Also, 90 percent confidence interval for RMSEA is 0.080 and 0.082.

The Normed Fit Index (NFI), The Relative Fit Index (RFI), The Incremental Fit Index (IFI) and the Comparative Fit Index (CFI) of this model are 0.984, 0.980, 0.984, and 0.980 respectively. All of these indices are indication of a good fitting model.

The expected cross-validation index (ECVI) is 0.413, which is very useful in the comparing different models and deciding how many factors is measured by SST mathematics sections. A model with the smallest ECVI value shows the greatest possibility to cross-validate.

In 2007 SST, three-factor model is plausible with having generally good fit index values as in 2006 SST. Therefore, three-factor model is selected as a good representation of what 2007 SST mathematics sections actually measures. This means that constructs being measured in 2006 SST mathematics sections cross validate with 2007 SST mathematics sections. This cross validation result across years is another validity evidence for three-factor structure of SST mathematics sections.

4.5 Relations among Mathematical Skills

The fifth research question asks “What mathematical skills are achieved at different ability levels of the students?”. In order to answer this question, by using Item Response Theory (IRT), item parameters of 2006 SST mathematics sections are estimated. These parameters are used to draw item map with respect to P50 and P80 response probabilities by using Wopgraph program, written by Cito. By investigating this bar graph, prerequisite nature of the mathematical skills can be identified. Then relation between mathematical skills and ability levels of students is investigated.

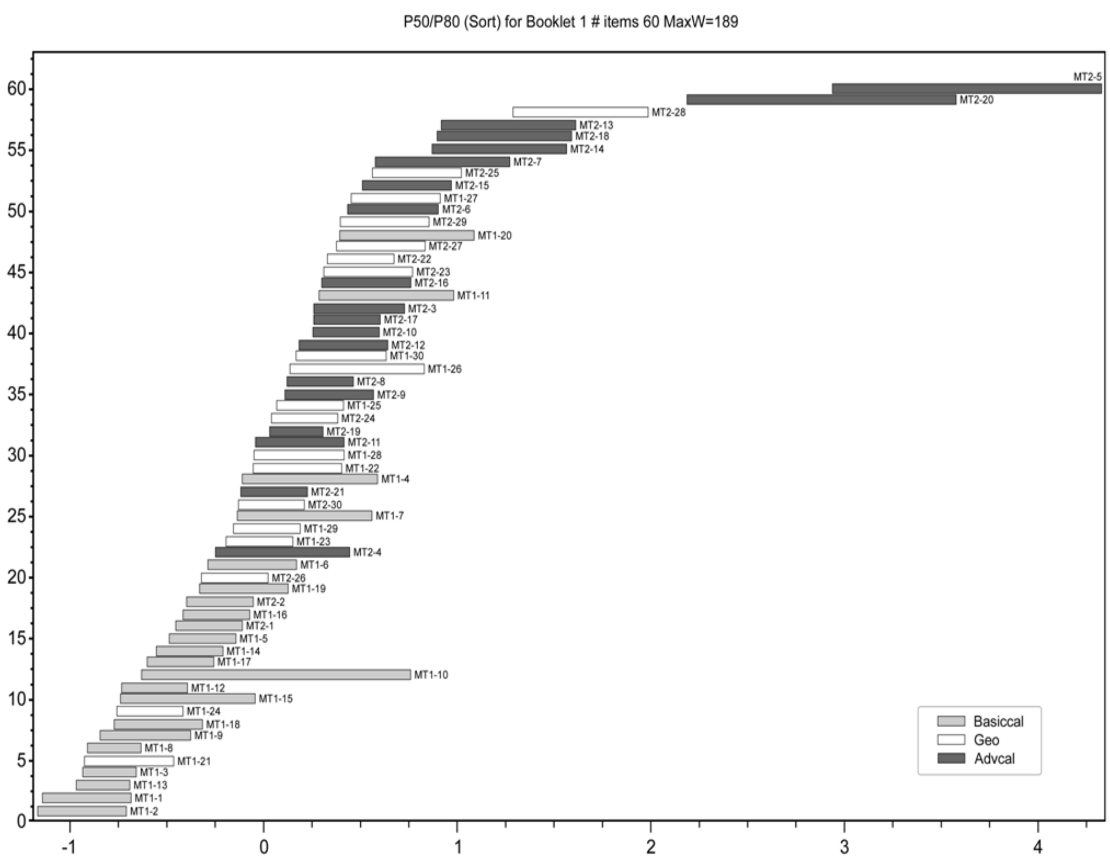
All mathematics items in 2006 SST are analyzed by IRT methods and item parameters are estimated. Appendix G gives the item discrimination index, the amount of student ability to have probability of 0.5 for a correct answer (P50), and the amount of student ability to have probability of 0.8 for a correct answer (P80) for each item. Also, in Figure 4.6, item map drawn according to P50 and P80 response probabilities are given. For an item, left side value of the bar represents P50 and right side value of the bar represents P80. Horizontal axis represents ability values of students.

Items represented in the bottom section are easy items and items represented in the upper section are difficult items. Therefore, item MT1.2 is the easiest item and item MT2.5 is the most difficult item. Besides that, a student who have a probability of 0.5 to correctly answer an item have higher probability to solve items below that item and have lower probability to solve items above that item. There is a hierarchical structure among these item bars according to IRT. By producing these bars, the main aim is to identify what mathematical skills are achieved at different ability levels of students. In order to identify this, item bars in Figure 4.6 are investigated whether similar items creates definable groups together.

Figure 4.6 indicate that items associated with *Basic Computation Ability* (Basiccal) grouped together at the bottom of the graph. Items associated with *Geometry Ability* (Geo) and items associated with *Advanced Computation Ability* (Advcal) grouped above *Basic Computation Ability* items. This means that lower ability level is necessary for a student to solve items in *Basic Computation Ability* than items in *Geometry Ability* and *Advanced Computation Ability*. Therefore, *Basic Computation Ability* is a prerequisite to be successful in these other abilities measured in SST. For example, according to Figure 4.6, a student with ability level of -0.25 has at least 0.5 probabilities to solve almost all *Basic Computation Ability* items, whereas, same student has less than 0.5 probabilities to solve almost all *Geometry Ability* items and *Advanced Computation Ability* items. According to this bar graph, students have higher probability to solve *Basic Computation Ability* items than *Geometry Ability* items and *Advanced Computation Ability* items.

However, it is seen that *Geometry Ability* items and *Advanced Computation Ability* items do not have separate group of bars. Therefore, there is not clear hierarchical relation between *Geometry Ability* and *Advanced Computation Ability*. For further analysis, this implied relation among mathematical skills will be tested by a structural equation model.

Figure 4.6 P50 and P80 response probabilities of Items



4.6 Structural Equation Models related to Mathematical Abilities

The sixth research question asks "What are the relationships among the dimensions defined in the mathematics subtests of the SST?". In order to answer this question, proposed hierarchical relation among mathematical skills by using a structural equation model is tested. In fifth research question it was proposed that *Basic Computation Ability* is a prerequisite to be successful in *Geometry Ability* and *Advanced Computation Ability*. Figure 4.7 shows the diagram for this relation. Fit indices values for this structural equation model are given below.

The chi-square, χ^2 = 35879.189, is significant with df=63, and the significance level is p=0.00. However, the χ^2 value is highly affected from sample size; as the sample size increases so does χ^2 (Schumacker & Lomax, 2004). Another criteria directly related with chi-square is the ratio of the chi-square value to the degrees of freedom which is 569.510 in this study. Although this value is very large than expected value of 2, it is impossible to have nonsignificant chi-square related values with sample size of 872 956.

The independence chi-square test value is 1762706.174 with degrees of freedom of 78 and this test is significant. Significant independence chi-square test means that there is some relationship among variables as desired.

The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI) of this model is 0.935 and 0.907 respectively. Since GFI value is close to 0.95 and AGFI is higher than 0.90, it is secure to say this model has a good fit to the data.

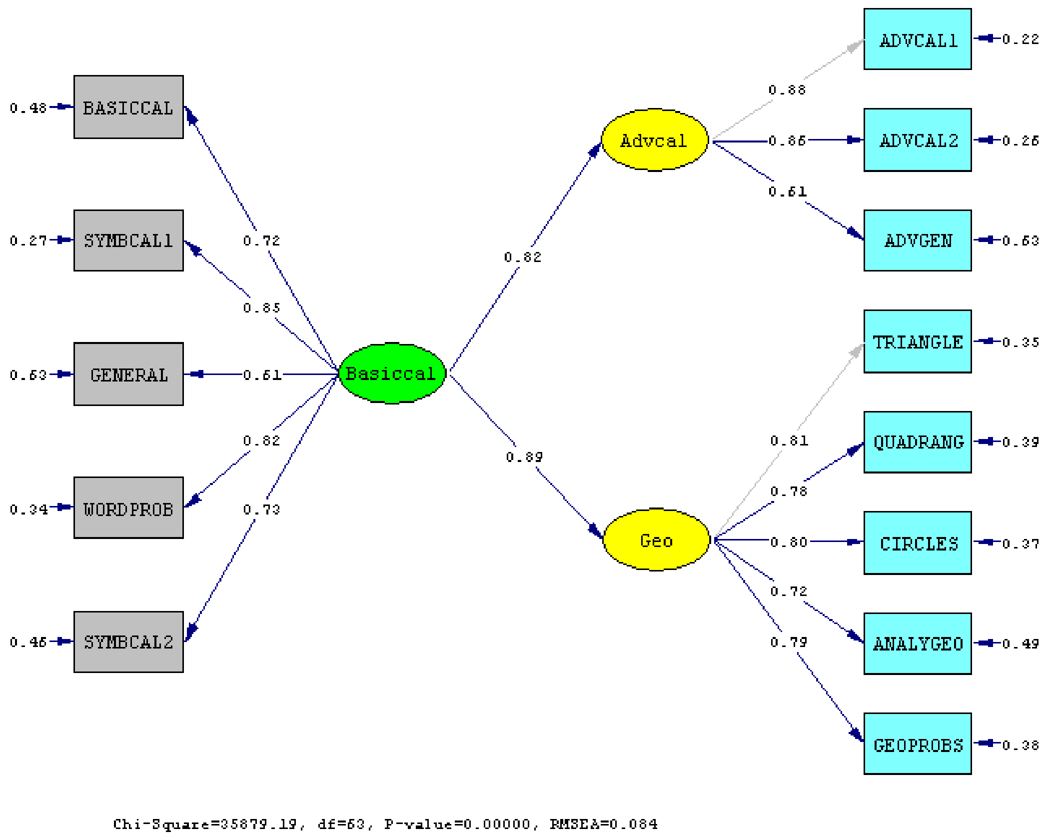
The Root-Mean-Square Residual (RMR) of this model is 0.099 which is higher than 0.05 and close to 0.10. This means for RMR, model has moderate fit. However, The Standardized Root Mean Square Residual is 0.049 which is less than

0.08, therefore indicating a good fit to the data. Also, the Root-Mean-Squared Error of Approximation (RMSEA) of the model is 0.084, higher than 0.06 and less than 0.10, shows a moderate fit to the data. Also, 90 percent confidence interval for RMSEA is 0.084 and 0.085.

The Normed Fit Index (NFI), The Relative Fit Index (RFI), The Incremental Fit Index (IFI) and the Comparative Fit Index (CFI) of this model are 0.981, 0.977, 0.981, and 0.981, respectively, which indicate good fit.

Generally, almost all of these indices are indication of a good fitting model. Therefore, the proposed relations among dimensions are supported by the data. It is accepted that *Basic Computation Ability* is a prerequisite to be successful in *Geometry Ability* and *Advanced Computation Ability*.

Figure 4.7 Structural Equation Model for Mathematical Abilities



In order to identify deeply relations among mathematical skills, relation of *Turkish Language Ability* of students with mathematical abilities is also investigated. A path analytical model in which Turkish subtest scores is added given in Figure 4.8. In this model it is proposed that *Turkish Language Ability* is related to *Basic Computation Ability* and *Basic Computation Ability* is a prerequisite to be successful in *Geometry Ability* and *Advanced Computation Ability*. Fit indices values for this structural equation model are given below.

The chi-square, χ^2 = 39725.323, is significant with df=75, and the significance level is p=0.00. However, the χ^2 value is highly affected from sample size; as the sample size increases so does χ^2 (Schumacker & Lomax, 2004). Another criteria directly related with chi-square is the ratio of the chi-square value to the degrees of freedom which is 529.670 in this study. Although this value is very large than expected value of 2, it is impossible to have nonsignificant chi-square related values with sample size of 872 956.

The independence chi-square test value is 1846748.173 with degrees of freedom of 91 and this test is significant. Significant independence chi-square test means that there is some relationship among variables as desired.

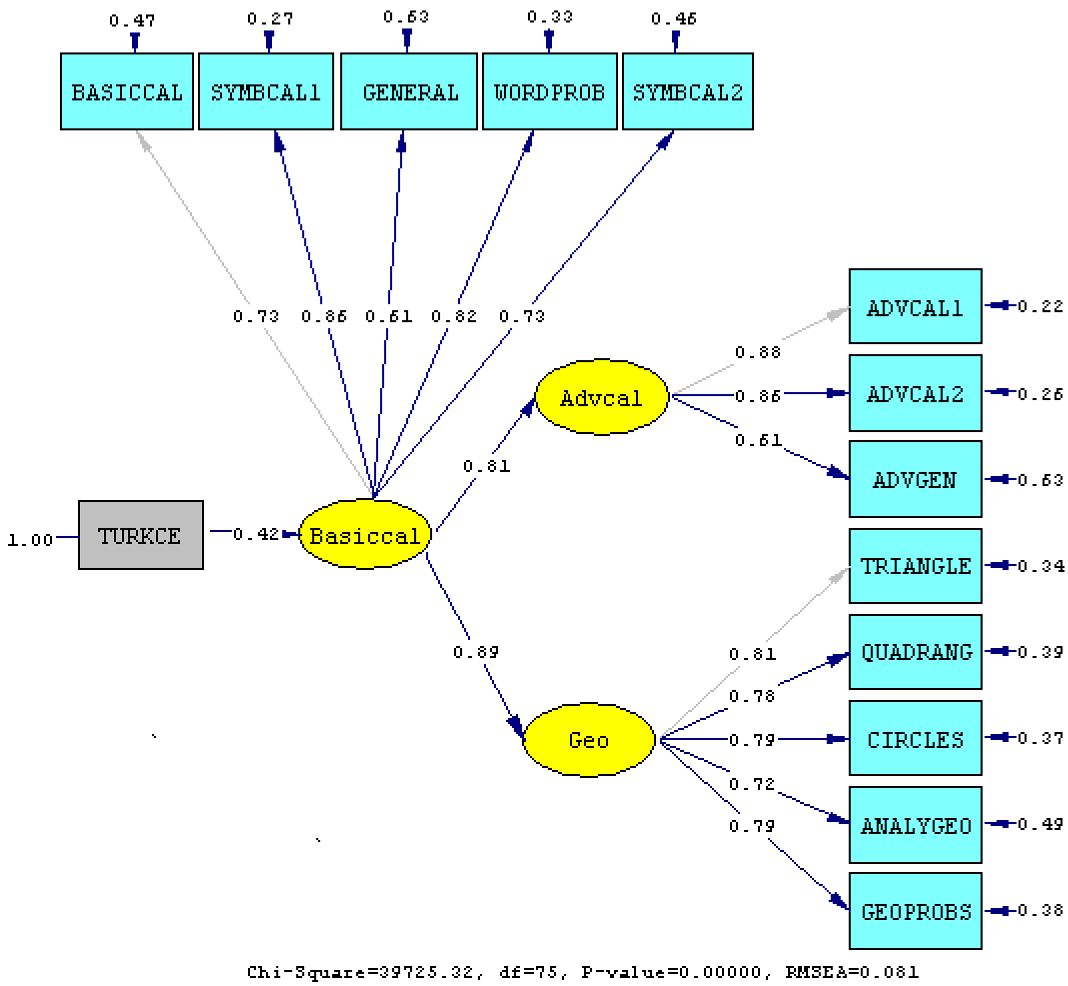
The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI) of this model is 0.934 and 0.907 respectively. Since GFI value is close to 0.95 and AGFI is higher than 0.90, it is secure to say this model has a good fit to the data.

The Root-Mean-Square Residual (RMR) of this model is 0.178 which is higher than 0.10. This means a poor fit to the data. However, The Standardized Root Mean Square Residual is 0.051 which is less than 0.08, therefore indicating a good fit to the data. Also, the Root-Mean-Squared Error of Approximation (RMSEA) of the model is 0.081, higher than 0.06, and less than 0.10 shows a moderate fit to the data. Also, 90 percent confidence interval for RMSEA is 0.081 and 0.082.

The Normed Fit Index (NFI), The Relative Fit Index (RFI), The Incremental Fit Index (IFI) and the Comparative Fit Index (CFI) of this model are 0.980, 0.976, 0.980, and 0.976, respectively, which indicate good fit.

Almost all of these indices are indication of an acceptable fitting model. Therefore, the proposed relations in the model are supported by the data. This means that *Turkish Language Ability* is related to *Basic Computation Ability* and *Basic Computation Ability* is a prerequisite to be successful in *Geometry Ability* and *Advanced Computation Ability*

Figure 4.8 Structural Equation Model for Mathematical Abilities and Turkish Language Ability



4.7 Testing Equality of Structural Equation across Groups

Seventh research question asks "Do the structural relationships defined in the mathematics subtest hold across gender and school type?". In order to investigate this question, equality of proposed structural equation model across groups is investigated. In this section equality of hierarchical relation among *Basic Computation Ability*, *Advanced Computation Ability*, and *Geometry Ability* across gender groups and school types are tested. Figure 4.7 gives the proposed relations among mathematical abilities.

4.7.1 Testing Equality of Structural Equation across Gender Groups

The equality of a structural equation model for males and females is tested in this section. Table 4.11 shows the fit indices values. With related to the invariance of the structural equation across males and females, the fit indexes for the individual samples and for the overall analysis is acceptable. This means that the proposed relations in the model supported by the data across gender groups. Therefore, the proposed structural equation model is invariant for males and females.

Table 4.11 Tests of Invariance of Structural Equation for Males and Females

	df	x^2	x^2/df	SRMR	GFI	CFI	NFI	RMSEA
Males				0.081	0.924			
Females				0.071	0.933			
Overall	154	80558.801	523.109			0.979	0.979	0.080

4.7.2 Testing Equality of Structural Equation across School Types

The equality of a structural equation for private and public schools is tested in this section. Table 4.12 shows the fit indices values. With related to the invariance of the structural equation across private and public schools, the fit indexes for the individual samples and for the overall analysis is acceptable. This means that the proposed relations in the model supported by the data across school types. Therefore, the proposed structural equation is invariant for males and females.

This cross validation result across gender groups and school types is another validity evidence for three-factor structure of SST mathematics sections and hierarchical relation among mathematical abilities.

Table 4.12 Tests of Invariance of Structural Equation for Private and Public Schools

	df	χ^2	χ^2/df	SRMR	GFI	CFI	NFI	RMSEA
Private				0.098	0.924			
Public				0.078	0.908			
Overall	154	94468.933	613.434			0.974	0.974	0.087

CHAPTER 5

CONCLUSION

In this chapter, results are summarized and discussed in seven main sections: (1) Cognitive Skills in 2006 SST, (2) Dimensions of Mathematics Subtests in 2006 SST, (3) Cross Validation of Results across Gender Groups and School Types, (4) Cross Validation of Results with 2007 SST, (5) Relations among Mathematical Skills, (6) Structural Equation Models related to Mathematical Abilities, (7) Testing Equality of Structural Equation across Groups. Limitations of the study, recommendations for future studies and suggestions to Student Selection and Placement Center are given at the end of the chapter.

5.1 Cognitive Skills in 2006 SST

The first research question asked "What cognitive skills are assessed in the mathematics subtest of the university entrance examination?". In order to answer this question content-wise evaluation of the test item content with respect to mathematical cognitive skills emphasized in the literature was used.

Content-wise evaluation, done by researcher and experts, showed that cognitive skills assessed in the mathematics subtest of the university entrance examination are not totally definable. Also, cognitive skills assessed by items in mathematics subtests do not match with theoretical framework of mathematics assessment emphasized in the literature. It is claimed both by researcher and experts that "Analysis/Reasoning" cognitive skill items are used very rarely in SST mathematics sections. This result is congruent with Berberoglu (1995) and Berberoglu et al. (1996) who stated that mathematics items in 1992 SST could be categorized only in comprehension level, none of the items could be categorized to measure application or analysis level defined by Bloom.

It is also important that there was not complete congruence among researcher and experts about classification of items. Therefore, in order to understand how items are grouped, whether item classifications are supported by the data, and what are the relations between cognitive skills, different approaches was used.

5.2 Dimensions of Mathematics Subtests in 2006 SST

The second research question asked "What are the dimensions of mathematics subtests (first and second section) in the student selection tests?". In order to answer this question and examine underlying constructs of mathematics sections, exploratory factor analysis was conducted; items were grouped; and confirmatory factor analysis was performed.

Results of exploratory factor analysis showed that there were five main factors in SST mathematics sections. In order to investigate these dimensions, items were grouped and five-factor model was tested by confirmatory factor analysis. However, fit indices of five-factor model were not supported by data. Therefore, four-factor, three-factor, two-factor and one-factor models which are formed according to relevant theory were tested. It is found that, four-factor model and three-factor model seem plausible with having good fit indices values. However, three-factor model had better fitting values than four-factor model. Two-factor model and one-factor model were not supported by data also.

Therefore, three-factor model is accepted as what SST mathematics sections measure. In accepted model, it was proposed that there are three correlated factors; one of them measures *Basic Computation Ability*, one of them measures *Advanced Computation Ability* and the other measures *Geometry Ability*. This result shows that SST is good in measuring "Computation/Knowing" cognitive process in which *Basic Computation Ability*, *Advanced Computation Ability* and *Geometry Ability are* mainly related to reaching successfully to a result by performing some steps of calculation.

On the other hand, in five-factor model it was proposed that *Generalization Ability* (related to "Comprehension/Knowing" cognitive process) and *Problem Solving Ability* (related to "Application/Analysis" cognitive process) were also measured in addition to three-factor model. Similarly, in four-factor model it was proposed that *Problem Solving Ability* (related to "Application/Analysis" cognitive process) was also measured in addition to three-factor model. However, these models were not accepted because of poor fit indices. Therefore, it is concluded that generalization and problem solving items are not functioning as it is aimed. These items measures some sort of computation ability. Consequently, these results showed that SST mathematics section do not measure beyond computation abilities.

A possible explanation of why *Problem Solving Ability* can not be differentiated from *Basic Computation Ability* in SST is follows: in all of problem solving items of SST, a short passage which defines a situation is given. After transferring this situation to a mathematical expression, performing some steps of calculations is expected. After successfully writing related mathematical expression, problem solving items in SST are nothing but conducting steps of basic calculation. This close relation of these items explains very high correlation between *Basic Computation Ability* and *Problem Solving Ability*. This outcome is congruent with the results of Tuna (1995) and Berberoglu (1996) who observed that computation and word problems generally grouped on a common factor whereas geometry items created different factor in 1993 SST and 1992 SST.

It is important to note that, SST should measure *Problem Solving Ability* as it is explained in literature. Conducting successfully some steps of calculations is not enough for today's world in respect to academic proficiencies and job market proficiencies. Items that measure *Problem Solving Ability* should be related to daily life situation; meaningful and concrete to students; and a decision making process should be a part of problem solving process. A situation different from above description, can not go beyond practice of algorithmic calculation (İş Güzel, 2009).

Problem Solving Ability items probably will function better with making some revision to these items. Sample problem solving items in Appendix E, Appendix F and a good example in SST can be an inspiration for these revisions. For instance, in 2006 SST, item MT1.18 is a good example for problem solving in which a decision making process by making some comparisons is involved. However, this item is only proper problem solving item in 2006 SST. For example, in 2006 SST, item MT1.16 is not a good example for problem solving. This item does not represent a situation people generally face in daily life situation and no decision making is required. Besides that, when item MT1.17 in 2006 SST and an item in PISA 2003 (Example 2 in Appendix E) compared, it is seen that both of them are related to travelling between cities. However, MT1.17 in SST does not contain any decision making process and situation in this item can not be encountered in real life, in contrast to item in PISA 2003.

Berberoglu (1995) explained that first part of 1992 SST mathematics section was consist of computation items, which are related to simple calculations and/or algorithms; word problem items, which are related to verbal problems in which it is required to understand the problem and carry out the related computations; and geometry items, in which it is required to apply basic concepts and principles of geometry. It is interesting to see that from 1992 to 2006 and 2007, in respect to skills and process measured in SST mathematics sections, nothing much is changed or improved.

2006 SST and 2007 SST mathematics sections have very high internal consistency values of 0.949 and 0.962 respectively. This result is consistent with the findings of study conducted by Uygun (2008) in which 2006 SST science items have internal consistency value of 0.94 and the findings of study conducted by Berberoglu et al. (1996) in which 1992 SST mathematics items and science items have internal consistency value of 0.93 and 0.94 respectively. Therefore, for overall reliability analysis, SST is a reliable instrument.

However, reliability analysis of items showed that some items in 2006 SST Mathematics sections have very low corrected item-total correlations. By eliminating these items from analyses, higher reliability values for overall SST and better fitting three-factor model is obtained. When the poor items are investigated, all of poor items belong to second mathematics section of SST.

It is very important to note that, in 2006, just after administration of SST, SSPC announced a press release. It is stated that item MT2.5 (Item MT2.15 according to items in SSPC website) had been investigated by subject area experts in universities and by experts of SSPC and it had been decided that item MT2.5 was not problematic as it was stated by press or sent objections (Student Selection and Placement Center, 2006c). However, this study showed that, in reliability analysis, this item is only item that had negative corrected item-total correlation among all items in 2006 SST.

Ideally, items have to be piloted (tested) before they are used. Poor items according to statistical analysis have to be identified, and they should be removed from test designs. Besides that, after administration of SST to students, post statistical analysis should be conducted in order to identify poor items again. Better tests and more valid results can be obtained if these procedures are followed. Inquiring items only subjectively by subject-area experts is not enough to have best SST. There are many examples of standard tests in which items are piloted before they are used and analyzed after the administrations for checking item statistics, like, Cito Turkey Pupil Monitoring System (ÖİS) and PISA.

5.3 Cross Validation of Results across Gender Groups and School Types

The third research question asked "Do dimensions of 2006 SST mathematics subtests provide stable factorial structure across gender groups and school types?". In order to answer this question, multiple group analysis related to accepted three-factor model for groups was conducted. Three-factor model was tested about the invariance of the factors across gender groups and school types.

For males and females, it is found that the number of factors is invariant; the factor correlations are invariant; error variances are invariant; and factor loadings are invariant. All of these models had signs of good fit. For private and public schools it is also found that the number of factors is invariant; the factor correlations are invariant; error variances are invariant; and factor loadings are invariant. All of these models had signs of good fit.

For multiple group analysis, invariant factor structure across gender groups and school types indicates that SST mathematics section is operating similarly for these groups. This cross validation result is another validity evidence for three-factor structure of SST mathematics sections.

This result is congruent with the results of Aslan (2000), who observed factorial structure of was not different between gender groups in 1998 SST and of

Tuna (1995), who observed the structure of 1993 SST content which is assessed by reliability estimates was equal across gender. It is important to note that, for these subgroups, structure of SST which is measured in terms of number of factors, factor correlations, error variances, and factor loadings are invariant as a overall analysis. However, item level analysis (like DIF) conducted previously in other studies showed differences for these subgroups.

5.4 Cross Validation of Results with 2007 SST

Fourth research question asked "Do dimensions of 2006 SST mathematics subtests provide stable factorial structure across years?". In order to answer this question, confirmatory factor analysis related to three-factor model for 2007 SST was reconducted.

Results from confirmatory factor analysis of 2007 SST is used to cross validate the results of what actually SST Mathematics sections measure and. Cross validation results also showed that three-factor model has better fit indices compared to other models. Therefore, three-factor model is accepted and cross validated as what SST mathematics sections measure. Therefore, more confidently it can be stated that SST mathematics sections only measures *Basic Computation Ability*, *Advanced Computation Ability* and *Geometry Ability*. By conducting similar analysis in two consecutive years, generalizability of the results is assessed. It can be stated that results obtained are similar and can be generalizable. This cross validation result is another validity evidence for three-factor structure of SST mathematics sections.

On the other hand, in 2006 and 2007 data, it was proposed that *Problem Solving Ability* is also measured as one of the main factors. However, this hypothesis is not accepted in both year analyses. Therefore, problem solving questions in SST are not functioning as it is intended, but measures computation ability as many of other SST questions. These results showed that SST mathematics sections for both years do not measure beyond computation abilities.

5.5 Relations among Mathematical Skills

The fifth research question asked "What mathematical skills are achieved at different ability levels of the students?". In order to answer this question, 2006 SST item parameters by using Item Response Theory (IRT) were estimated. With using results of this analysis, item map with respect to P50 and P80 response probabilities is drawn.

It is found that items associated with *Basic Computation Ability* (Basiccal) grouped together at the bottom of the bars. Items associated with *Geometry Ability* (Geo) and items associated with *Advanced Computation Ability* (Advcal) grouped above *Basic Computation Ability* items. This means that lower ability level is necessary to solve items in *Basic Computation Ability* than items in *Geometry Ability* and *Advanced Computation Ability*. Therefore, *Basic Computation Ability* is a prerequisite to be successful in these other abilities measured in SST. However, it is seen that *Geometry Ability* items and *Advanced Computation Ability* do not have separate group of bars. Therefore, there is not clear relation between *Geometry Ability* and *Advanced Computation Ability*.

This relation implies that teaching strategies should be designed in a way that students should gain *Basic Computation Ability* before they start to learn other abilities defined in this study. Emphasis on teaching should be given to this hierarchical structure. *Geometry Ability* and *Advanced Computation Ability* are more complex abilities, and they are less likely to develop unless *Basic Computation Ability* is acquired.

5.6 Structural Equation Models related to Mathematical Abilities

The sixth research question asked “What are the relationships among the dimensions defined in the mathematics subtests of the SST?”. In order to answer this question, relation among mathematical skills by using a structural equation model was tested.

It is found that the proposed relation among mathematical abilities is supported by the data. Therefore, *Basic Computation Ability* is a prerequisite to be successful in *Geometry Ability* and *Advanced Computation Ability*. Besides that in order to identify deeply relations between mathematical skills, relation of *Turkish Language Ability* of students with mathematical abilities is also investigated. In this model it is proposed that Turkish language ability is related to *Basic Computation Ability* and *Basic Computation Ability* is a prerequisite to be successful in *Geometry Ability* and *Advanced Computation Ability*. It is found that the proposed relations are also supported by the data. Therefore, there is a significant relation between *Turkish Language Ability* and *Basic Computation Ability*.

5.7 Testing Equality of Structural Equation across Groups

Seventh research question asked “Do the structural relationships defined in the mathematics subtest hold across gender and school type?”. In order to investigate this question, equality of proposed structural equation model across gender groups and school type was investigated.

It is found that *Basic Computation Ability* is a prerequisite to be successful in *Geometry Ability* and *Advanced Computation Ability* for males and females and for private and public schools. Therefore, invariant structural equation model across gender groups and school types indicates that SST mathematics section is operating similarly for these groups.

5.8 Limitations of the Study

The main limitation of this study is that opinions of more experts, including item writers in SSPC, could be used to classify the items according to a guideline. Item grouping differences between these experts could be discussed together to have a congruence between them.

5.9 Recommendations for Future Studies

In 2010, new SST and university entrance system will be used for admission to universities. With using this new 2010 SST mathematics sections data, models tested in this study can be reconducted to see similarities and differences between 2010 SST, 2006 SST and 2007 SST results.

Besides that, the relations between identified factors in mathematics and other subject areas can also be analyzed. The higher order models for overall SST using other subtests can be compared and tested for future studies.

Further researchers can also investigate relations between identified dimensions in mathematics and student questionnaire, which is administered online by Research and Development department of SSPC. For example, effects of having different study habits on computation ability are worth to investigate.

Also, relations between dimensions that are measured in SST and other variables like future university graduation point average (GPA) of students or graduation school examination (ALES) results of students can be investigated. Investigating which factors are best predictors of these future variables are important for validation of results in different perspective.

5.10 Suggestions to Student Selection and Placement Center

Several suggestions about SST to SSPC are listed in this section. First of all, a manual or assessment framework which describes measurement properties and cognitive processes measured in SST should be published. It is important to note that for each item there should be only one definable cognitive process. Therefore, difference between a student who answered correctly similar types of items and a student who can not answered those items in terms of cognitive perspective will be clear (Berberoglu, 2009). After administering SST, whether proposed test plan is achieved should be analyzed by SSPC and controlled by item writers. If some of the aimed dimensions are not measured as it is planned, some revisions should be made. With these revisions done each year by using empirical results, it is believed that better SST will be developed.

Also, as it is explained, problem solving ability is not measured properly. Problem solving items are very close to computation ability items. They should be revised generally and what problem solving is should be congruent with the literature. Number of items measuring Computation Ability should be decreased and items measuring application and analysis dimension of cognitive processes should be used and increased.

SSPC should use items that measure beyond pure computation ability of students. Higher order thinking abilities of students should be measured also. In order to measure these types of abilities, not only multiple choice items, but also various item formats like open-ended items can be used in test plans. Besides that, almost all items should represent situations that students may encounter in their daily life and these situations should be meaningful to them.

One of the objectives of measurement and evaluation is to improve quality of education by giving feedback to students, and to institutions. SST only reports norm referenced results. Therefore, there is not any information about what is difference between two students one is accepted to a department and one is not accepted to same department in terms of mathematical abilities (Berberoglu, 2009). Besides norm references results, SSPC should report criterion referenced results, in which, abilities and proficiency levels of students related to each subtest can be identified. Therefore, besides students, universities and departments will have more detailed information about students admitted to their departments in terms of cognitive characteristics of students. For example, Cito Turkey Pupil Monitoring System (ÖİS) and PISA report criterion referenced results by giving proficiency level descriptions for each student.

Ideally, items should be piloted (tested) before they are used. Therefore, it can be evaluated whether items are appropriate to students or not (Berberoglu, 2009). Problematic items (poor items, items that contain DIF, etc.) according to statistical analysis should be identified, and they should be removed from test designs. Besides that, after administration of SST to students, post statistical analysis should be conducted in order to identify problematic items again. Better test results can be obtained if these procedures are followed. Inquiring items only subjectively by subject-area experts is not enough to have best SST. In all empirical studies done about items of SST, (Berberoglu, 1995; Ogretmen, 1995; Yenal, 1995; Kalaycioglu, 2008) problematic items are found which shows there is a need to analyze items before and after they used.

By using the results of studies conducted previously, findings of this study and suggested other studies that will be conducted about SST, it is aimed that better SST can be developed and administered. It is expected that SST will have better psychometric properties, what is measured by SST will be clear and congruent with today’s world, and SST will give feedbacks to stakeholders.

REFERENCES

Arvey, R. D. (1992). Constructs and Construct Validation: Definitions and Issues. *Human Performance,* 5(1&2), 59-69.

Aslan, Ö. (2000). The Construct validity of the 1998 secondary school institutions student selection and placement test. Master thesis, Middle East Technical University, Turkey.

Berberoğlu, G. (1995). Differential Item Functioning (DIF) Analysis of Computation, Word Problem and Geometry Questions Across Gender and SES Groups. *Studies in Educational Evaluation,* 21, 439-456.

Berberoğlu, G. (1996). The University Entrance Examination in Turkey. *Studies in Educational Evaluation,* 22(4), 363-373.

Berberoğlu, G. (2009). Madde haritalama yöntemi ve Cito Türkiye Öğrenci İzleme Sistemi (ÖİS) uygulamalarında yeterlik düzeylerinin belirlenmesi. *Cito Eğitim: Kuram ve Uygulama, 3*, 13-24.

Berberoğlu, G., Dochy, F. J. R. C. , Moerkerke, G. (1996). Psychometric Evaluation of Entry Assessment in Higher Education: A Case Study. *European Journal of Psychology of Education,* Vol. XI, No. 1, 25-41.

Bloom, B. S., Hastings, J. T., Madaus, G. F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. McGraw-Hill Book Company

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 137–162). Newbury Park, CA: Sage.

Butler, S. F.; Fernandez, K.; Christine, B.; Budman, S. H.; Jamison R. N. (2008). Validation of the Revised Screener and Opioid Assessment for Patients with Pain (SOAPP-R). *J Pain*, 9(4), 360–372.

Blunch, N. J. (2008). *Introduction to Structural Equation Modelling using SPSS and AMOS.* London: Sage Publications.

Byrne, B. M. (1998). *Structural Equation Modeling with LISREL, PRELIS , and SIMPLIS: Basic Concepts, Applications, and Programming.* New Jersey: Lawrence Erlbaum Associates, Inc.

Byrnes, J. P. & Takahari, S. (1993). Explaining Gender Difference on SAT-Math Items. *Developmental Psychology*, 29(5), 805-810

Can, S. (2003). The Analyses of secondary education institutions student selection and placement tests verbal section whith respect to item response theory models. Master thesis, Middle East Technical University, Turkey.

Carlstedt, B. & Gustafsson, J. (2005). Construct validation of the Swedish scholastic aptitude test by means of the Swedish enlistment battery. *Scandinavian Journal of Psychology*. 46, 31-42.

Chen, Y. & Thompson, M. S. (2004). Confirmatory factor analysis of a school self-concept inventory. Paper presented at the annual conference of Arizona Educational Research Organizaiton April 12-14 Tempe, Arizona, USA.

College Board, http://www.collegeboard.com/press/releases/150054.html, last accessed date: 03-11-2006

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443-507).

Çınar, S. (2006). Üniversiteye giriş sınavının üniversite sınavına giren öğrenciler üzerindeki etkileri. Master thesis, Fırat University, Turkey

Dayıoğlu, M. & Aşık, S. T. (2004). *Gender Differences in Academic Performance in a Large Public University in Turkey.* ERC Working Papers in Economics, 04/17.

DeVellis, R. F. (2003). *Scale Development: theory and application.* California: Sage Publications.

Fraenkel, J. R. & Wallen, N. E. (2005). *How To Design And Evaluate Research In Education.* New York: McGraw-Hill.

George, D. & Mallery, P. (2001). *SPSS for Windows Step by Step: A Simple Guide and Reference.* Boston: Allyn & Bacon.

Gregory, K., & Clarke, M. (2003). High-Stakes assessment in England and Singapore. *Theory Into Practice*, *42*(1), 66-74.

Hu, L.-T. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives, *Structural Equation Modeling,* 6, 1–55.

İş Güzel, Ç. (2009). Cito Türkiye Öğrenci İzleme Sistemi'nde (ÖİS) Problem Çözme Becerilerinin Ölçülmesi Etkinlik ve Sorularla Örneklendirilmesi ve Değerlendirilmesi. *Cito Eğitim: Kuram ve Uygulama,* 4, 10-20.

Jöreskog, K. G., & Sörbom, D. (1993). LISREL 8: *Structural equation modeling with the SIMPLIS command language.* Chicago: Scientific Software International.

Kalaycıoğlu, D. B. (2008). Öğrenci Seçme Sınavı'nın Madde Yanlılığı Açısından İncelenmesi. PHD thesis, Hacettepe University, Turkey.

Kline, R.B. (1998). *Principles and Practice of Structural Equation Modeling*, The Guilford Press, New York.

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, Middle East Technical University, Turkey.

Kollu, S. (2006). The effects of private schools on public school achievement. Master thesis, Bilkent University, Turkey.

Köksal, E. A. (2002). The Assessment of the biology items in the 1998-2001 secondary school institutions student selection and placement tests. Master thesis, Middle East Technical University, Turkey.

Krulik, S. & Rudnick, J. A. (1989). Problem solving: A handbook for senior high school teachers. Upper Saddle River, NJ: Allyn and Bacon.

Mau, W. & Lynn, R. (2001). Gender Differences on the Scholastic Aptitude Test, the American College Test and College Grades. *Educational Psychology*, 21(2), 133-136.

MacCallum, R. C.; Browne, M. W.; Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130-149.

McDonald, R. P., & Moon-Ho, R. H. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, *7*(1), 64-82.

Mullis, I. V. S.; Martin, M. O.; Ruddock, G. J.; O'Sullivan, C. Y.; Arora, A.; Erberber, E. (2007). TIMSS 2007 Assessment Frameworks.

Noddings, N. (1985). Small groups as a setting for research on mathematical problem solving. In E. A. Silver (Ed). *Teaching and learning mathematical problem solving:*

Multiple research perspectives. (pp. 345-359). New Jersey: Lawrence Erlbaum Associates.

OECD (2003). The PISA 2003 assessment framework – Mathematics, reading, science and problem solving knowledge and skills.

OECD (2004). Problem solving for tomorrow's world – first measures of cross-curricular competencies from PISA 2003.

Öğretmen, T. (1995). Differential item functioning analysis of the verbal analysis of the verbal ability section of the first stage of the university entrance examination in Turkey. Master thesis, Middle East Technical University, Turkey.

Pallant J. (2007). *SPSS survival manual: a step by step guide to data analysis using SPSS*. Allen & Unwin.

Rubinstein, M. F. (1980). A decade of experience in teaching an interdisciplinary problem-solving course. In D. T. Tuma, & F. Reif (Eds). *Problem solving and education: Issues in teaching and research* (pp. 25-38). New Jersey: Lawrence Erlbaum Associates.

Schumacker, R. E. & Lomax, R. G. (2004). *A Beginners Guide to Structural Equation Modeling*. New Jersey: Lawrence Erlbaum Associates, Inc.

Schwieger, R. D. (1999). Teaching mathematical problem solving. In *Teaching elementary school mathematics*. (pp. 112-144). Canada: Wadsworth Publishing Company.

Segars, A. H. & Grover, V. (1993). Re-examining perceived ease of use and usefulness: A confirmatory factor analysis. *MIS Quarterly*, 17(4), 517-525.

Sharma, S. (1996). Applied multivariate techniques. New York: J. Wiley.

Stricker, L. J., Rock, D. A., Lee, Y.-W. (2005). *Factor structure of the LanguEdge test across language groups* (TOEFL Monograph Series Rep. No. MS-32: ETS Research Rep. No. 05-12). Princeton, NJ: ETS.

Stricker, L. J., & Rock, D. A. (2008). *Factor structure of the TOEFL Internet-Based test across subgroups* (TOEFL iBT Research Report Rep. No. TOEFLiBT-07: ETS Research Rep. No. 08-66). Princeton, NJ: ETS.

Student Selection and Placement Center (2006a), http://www.osym.gov.tr/BelgeGoster.aspx?F6E10F8892433CFF7A2395174CFB32E1782D5F4001E6D616, last accessed date: 10-11-2006

Student Selection and Placement Center (2006b), http://www.osym.gov.tr/BelgeGoster.aspx?F6E10F8892433CFF7A2395174CFB32E15F640FC6104C033D, last accessed date: 10-11-2006

Student Selection and Placement Center (2006c), http://www.osym.gov.tr/Genel/BelgeGoster.aspx?F6E10F8892433CFF1A9547B61DAFFE2A9DB0AA138C133A8F, last accessed date: 19-01-2010

Student Selection and Placement Center (2006d), ftp://dokuman.osym.gov.tr/OSYS2006KILAVUZ/10_BOLUM7.pdf, last accessed date: 23-10-2009

Student Selection and Placement Center (2006e), ftp://dokuman.osym.gov.tr/OSYS2006KILAVUZ/6_BOLUM3.pdf, last accessed date: 20-01-2010

Student Selection and Placement Center (2009a), http://www.osym.gov.tr/Genel/BelgeGoster.aspx?F6E10F8892433CFFD4AF1EF75F7A796823BE6F87556AA8A4, last accessed date: 18-01-2010

Tabachnick, B. & Fidell, L. S. *Using multivariate statistics.* Boston: Allyn &Bacon

Tansel, A. (1999). Türkiye'de ve Seçilmiş Ülkelerde Eğitimin Getirisi. *ODTÜ Gelişme Dergisi*, 26(3-4): 453-472

Tuna, S. (1995). A Factorial analysis of the First Stage of the University Entrance Examination. Master thesis, Middle East Technical University, Turkey.

Treat, A., Weersing, V. R. Clinical psychology. In B. S. Everitt & D. C. Howell (Eds), *Encyclopedia of statistics in behavioral science* (pp. 290-301). John Wiley & Sons.

Ullman, J. B. (2001). Structural equation modeling. In B. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (4^{th} ed., pp.653-771). Boston: Allyn &Bacon

Uygun, N. (2008). Validity of science items in the student selection test in Turkey. Master thesis, Middle East Technical University, Turkey.

Vernille, Kristy (2001), http://www.math.umd.edu/~dac/650/vernillepaper.html, last accessed date: 18-01-2010

Yenal, E. (1995). Differantial item functioning analysis of the quantitative ability section of the first stage of the university entrance examination in Turkey. Master thesis, Middle East Technical University, Turkey.

World Almanac & Book of Facts (2005). *SAT Mean Verbal and Math Scores of College-Bound Seniors, 1975-2004.*

Zumbo, B. D. (2005). Structural equation modeling and test validation. In B. S. Everitt & D. C. Howell (Eds), *Encyclopedia of statistics in behavioral science* (pp. 1951-1958). John Wiley & Sons.

APPENDIX A

2006 SST MATHEMATICS SECTION I ITEMS

MT1.1 $5-(-2+3)$ **işleminin sonucu kaçtır?** A) -1 B) 0 C) 4 D) 6 E) 10	MT1.2 $\dfrac{1-\frac{1}{2}+\frac{1}{3}}{\frac{1}{2}-\frac{1}{3}+\frac{1}{4}}$ **işleminin sonucu kaçtır?** A) 2 B) 1 C) 0 D) -1 E) -2
MT1.3 $\dfrac{\frac{1}{0,1}+\frac{0,2}{0,02}}{\frac{0,3}{3}}$ **işleminin sonucu kaçtır?** A) 0,2 B) 0,3 C) 20 D) 30 E) 200	MT1.4 a, b ve c tamsayıları için $a>b>c>0$ ve $c=a-b$ dir. **a ve b nin en büyük ortak böleni 4 olduğuna göre, aşağıdakilerden hangisi <u>kesinlikle</u> çift sayıdır?** A) $\frac{a+b}{4}$ B) $\frac{b+c}{4}$ C) $\frac{a}{4}+c$ D) $\frac{a-c}{4}$ E) $\frac{a+b+c}{4}$
MT1.5 **$0<x<1$ olmak üzere,** $a=x$ $b=x^2$ $c=\frac{1}{\sqrt{x}}$ **olduğuna göre, aşağıdaki sıralamalardan hangisi doğrudur?** A) $a<b<c$ B) $b<a<c$ C) $b<c<a$ D) $c<a<b$ E) $c<b<a$	MT1.6 $4^{2m-1}>\left(\frac{1}{16}\right)^{m+7}$ **eşitsizliğini sağlayan <u>en küçük</u> m tamsayısı kaçtır?** A) -4 B) -3 C) -2 D) 1 E) 2

MT1.7

$x = |\sqrt{5} - 3|$

$y = |x - 5|$

$z = |y - 2|$

olduğuna göre, z kaçtır?

A) $\sqrt{5}$ B) $2+\sqrt{5}$ C) $4+\sqrt{5}$

D) $10-\sqrt{5}$ E) $5-\sqrt{5}$

MT1.8

Hangi sayının 3 eksiğinin $\frac{2}{3}$ ü aynı sayının 5 eksiğine eşittir?

A) 6 B) 7 C) 8 D) 9 E) 12

MT1.9

Üç basamaklı 82A sayısının 9 ile bölümünden elde edilen kalan 7 ve üç basamaklı 3AB sayısının 9 ile bölümünden elde edilen kalan 2 dir.

Buna göre, üç basamaklı BAA sayısının 9 ile bölümünden elde edilen kalan kaçtır?

A) 3 B) 4 C) 5 D) 6 E) 7

MT1.10

5 e tam olarak <u>bölünemeyen</u> pozitif tamsayılar küçükten büyüğe doğru sıralanıyor.

Bu sıralamadaki 100. sayı aşağıdakilerden hangisidir?

A) 120 B) 124 C) 130 D) 134 E) 140

MT1.11

a ve b sıfırdan farklı gerçel sayılar olmak üzere,

$a * b = \frac{a}{b} = a - b$

olduğuna göre, a + b toplamı kaçtır?

A) $\frac{-3}{2}$ B) $\frac{-3}{4}$ C) 0

D) $\frac{1}{2}$ E) $\frac{2}{3}$

MT1.12

$\frac{1}{a^2} + \frac{4}{a} + 4 = 0$

olduğuna göre, a kaçtır?

A) $\frac{1}{2}$ B) 1 C) -2 D) -1 E) $\frac{-1}{2}$

MT1.13

a pozitif bir gerçel sayı ve

$a^4 - 2a^2 = 8$

olduğuna göre, a kaçtır?

A) $\frac{1}{8}$ B) $\frac{1}{4}$ C) $\frac{1}{2}$ D) 1 E) 2

MT1.14

$$\frac{3^{20} - 3^{10}}{(3^5+1)(3^5-1)}$$

işleminin sonucu kaçtır?

A) 1 B) 9 C) 3^5 D) 3^{10} E) 3^{15}

MT1.15

Aynı evde oturan bir grup arkadaş ev kirasını eşit olarak paylaşıyor.

Eve yeni bir arkadaş gelince kira için kişi başına düşen para % 20 azaldığına göre, <u>yeni arkadaşın gelmesiyle</u> evde oturan kişi sayısı kaç olmuştur?

A) 3 B) 5 C) 6 D) 8 E) 9

MT1.16

Ahmet parasının $\frac{2}{3}$ ü ile 3 gömlek ve 2 kravat, kalan parasıyla da 1 gömlek ve 3 kravat alabiliyor.

Buna göre, bir gömleğin fiyatı bir kravatın fiyatının kaç katıdır?

A) 2 B) 3 C) 4 D) 5 E) 6

MT1.17	MT1.18

MT1.17

Bir araç, iki kent arasındaki yolu saatte ortalama 60 km hızla gidip, hiç mola vermeden saatte ortalama 80 km hızla dönerek yolculuğu 7 saatte tamamlıyor.

Bu iki kent arasındaki uzaklık kaç km dir?

A) 240 B) 280 C) 300 D) 320 E) 360

MT1.18

Bir mağazada pantolon p, kazak k, tişört t YTL den satılmaktadır. Aşağıdaki tabloda Defne, Engin ve Mutlu'nun bu mağazadan aldıkları pantolon, kazak, tişört sayıları gösterilmiştir.

	Aldığı Giysi Sayısı		
	Pantolon	Kazak	Tişört
Defne	2	1	1
Engin	1	1	2
Mutlu	1	2	1

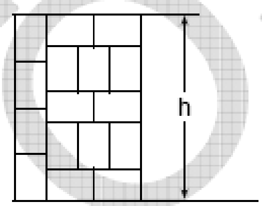
Aldıkları giysiler için en az parayı Engin, en çok parayı Mutlu ödediğine göre, aşağıdaki sıralamalardan hangisi doğrudur?

A) $p<t<k$ B) $k<t<p$ C) $k<p<t$

D) $t<p<k$ E) $t<k<p$

MT1.19

Aşağıdaki şekil, eş tuğlaların yatay ve dikey döşenmesiyle oluşturulan bahçe duvarının bir bölümünü göstermektedir.



Tuğlaların ayrıtlarının uzunlukları cm cinsinden <u>birer tamsayı</u> olduğuna göre, duvarın h ile gösterilen yüksekliği kaç cm <u>olabilir</u>?

A) 90 B) 100 C) 120 D) 140 E) 150

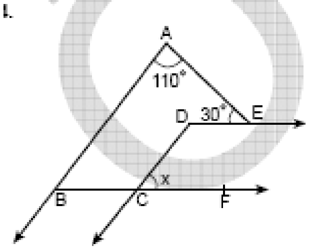
MT1.20

A = { 1, 2, 3, 4 } kümesinin elemanlarıyla, <u>en az</u> iki basamağındaki rakamı aynı olan <u>üç basamaklı</u> kaç sayı yazılabilir?

A) 52 B) 40 C) 38 D) 30 E) 24

MT1.21

I.



AB // DC

DE // CF

$m(\widehat{BAE}) = 110°$

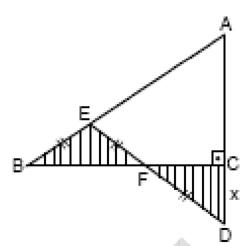
$m(\widehat{AED}) = 30°$

$m(\widehat{DCF}) = x$

Yukarıdaki verilere göre, x kaç derecedir?

A) 40 B) 50 C) 60 D) 70 E) 80

MT1.22



ABC bir üçgen

$BC \perp AD$

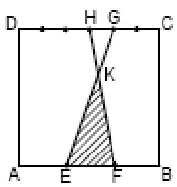
$|BE| = |EF| = |FD|$

$|CD| = x$

Şekildeki taralı bölgelerin alanları toplamı 12 cm² ve $|BC| = 8$ cm olduğuna göre, x kaç cm dir?

A) $\sqrt{2}$ B) $\sqrt{3}$ C) 2 D) 3 E) 4

MT1.23

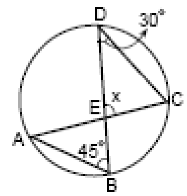


Şekildeki ABCD karesinin [AB] kenarı 3 eş parçaya, [CD] kenarı da 6 eş parçaya bölünmüştür.

[GE] ve [HF] doğru parçaları yardımıyla oluşturulan KEF üçgeninin alanı 4 cm^2 olduğuna göre, |AB| uzunluğu kaç cm dir?

A) 12 B) 9 C) 8 D) 6 E) 3

MT1.24



$m(\widehat{BDC}) = 30°$

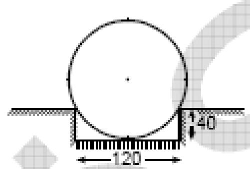
$m(\widehat{ABD}) = 45°$

$m(\widehat{DEC}) = x$

Yukarıdaki verilere göre, x kaç derecedir?

A) 95 B) 100 C) 105 D) 110 E) 115

MT1.25

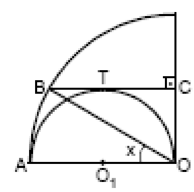


Dikey kesiti çember biçiminde olan bir iş makinesi lastiği; derinliği 40 cm, boyu 120 cm, dikey kesiti dikdörtgen biçiminde olacak şekilde oyulmuş bir altlığa şekildeki gibi tam oturtularak sergilenmektedir.

Buna göre, lastiğin dikey kesitinin yarıçapı kaç cm dir?

A) 75 B) 72,5 C) 70 D) 67,5 E) 65

MT1.26



$BC \perp OC$

$AO \perp OC$

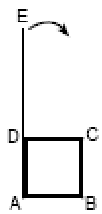
$m(\widehat{AOB}) = x$

Şekildeki O_1 merkezli yarım çember, O merkezli çeyrek çembere A noktasında, [BC] doğru parçasına da T noktasında teğettir.

Buna göre, x kaç derecedir?

A) 15 B) 20 C) 30 D) 45 E) 60

MT1.27

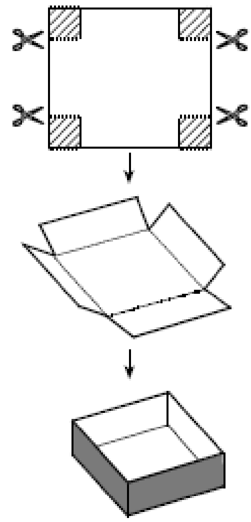


Şekilde verilen 8 cm uzunluğundaki DE ipi, gergin durumda tutularak, çevre uzunluğu 8 cm olan ABCD karesi biçimindeki çerçevenin etrafına saat yönünde döndürülerek sarılıyor.

İpin E ucu karenin D köşesine geldiğinde ipin taradığı alan kaç cm^2 olur?

A) 20π B) 22π C) 24π D) 28π E) 30π

MT1.28

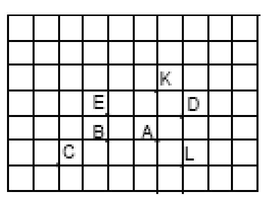


Bir kenar uzunluğu 16 cm olan kare şeklindeki kartonun köşelerinden bir kenar uzunluğu 3 cm olan birer kare kesilerek çıkartılıyor ve kalan karton parçası kıvrılarak şekildeki gibi üstü açık bir kutu yapılıyor.

Bu kutunun hacmi kaç cm^3 tür?

A) 200 B) 240 C) 250 D) 300 E) 360

MT1.29



Birim karelere bölünmüş bir kâğıt üzerinde A, B, C, D, E, K, L noktaları şekildeki gibi işaretlenmiştir. Bu kareli kâğıda A, B, C, D, E noktalarından biri orijin olacak biçimde bir dik koordinat sistemi yerleştiriliyor.

K ve L noktalarının orijine uzaklıkları eşit olduğuna göre, orijin aşağıdakilerden hangisidir?

A) A B) B C) C D) D E) E

MT1.30

A(−3, 4) noktasının $y = -x$ doğrusuna göre simetriği B ve B nin Ox eksenine göre simetriği C ise |BC| uzunluğu kaç birimdir?

A) $\frac{9}{2}$ B) $\frac{7}{2}$ C) 8 D) 6 E) 5

APPENDIX B

2006 SST MATHEMATICS SECTION II ITEMS

MT2.1

$$\left(\frac{x}{1+x}-\frac{1}{1-x}\right):\left(\frac{1}{1+x}+\frac{x}{1-x}\right)$$

işleminin sonucu aşağıdakilerden hangisidir?

A) 1 B) -1 C) x D) $1-x$ E) $1+x$

MT2.2

$$\frac{y^3+27}{y^2-2y-3}\cdot\frac{(y-3)\left(y^2-1\right)}{y^2-3y+9}$$

ifadesinin sadeleştirilmiş biçimi aşağıdakilerden hangisidir?

A) $(y+3)(y-1)$ B) $(y+3)(y-2)$

C) $(y+1)(y-3)$ D) $(y-1)(y-2)$

E) $(y-1)(y-3)$

MT2.3

$$|z|+z=3-2i$$

eşitliğini sağlayan z karmaşık sayısı aşağıdakilerden hangisidir?

A) $\frac{3}{5}-2i$ B) $\frac{5}{6}-2i$ C) $\frac{3}{4}+2i$

D) $\frac{2}{3}-3i$ E) $\frac{3}{5}+3i$

MT2.4

Aşağıdaki tabloyla değişmeli olmayan $(G,*)$ grubu verilmiştir (Örneğin, bu grupta $c*d=e$, $d*c=f$ dir.).

*	a	b	c	d	e	f
a	a	b	c	d	e	f
b	b	c	a	f	d	e
c	c	a	b	e	f	d
d	d	e	f	a	b	c
e	e	f	d	c	a	b
f	f	d	e	b	c	a

Buna göre, $b*(x*c)=d$ eşitliğini sağlayan x elemanı aşağıdakilerden hangisidir?

A) f B) e C) d D) c E) b

MT2.5

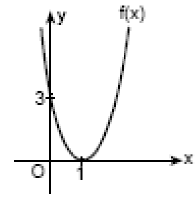
A boş olmayan bir küme olmak üzere, A dan A ya f ve g fonksiyonları tanımlanmıştır.

$(fog)(x) = f(g(x))$ ile verilen fog bileşke fonksiyonu <u>bire bir</u> ise aşağıdakilerden hangisi <u>kesinlikle</u> doğrudur?

A) f örtendir. B) g örtendir. C) f bire birdir.

D) g bire birdir. E) gof bire birdir.

MT2.6



$f(x)$ fonksiyonunun grafiği, şekildeki gibi, Ox eksenine $(1, 0)$ noktasında teğet olan ve $(0, 3)$ noktasından geçen paraboldür.

Buna göre, $f(3)$ kaçtır?

A) 3 B) 4 C) 6 D) 7 E) 12

MT2.7

$$(1-m)x^2 + 4x + m^2 - 4 = 0$$

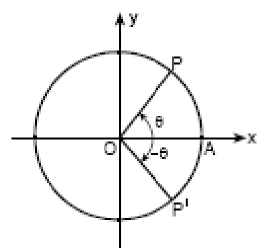
denkleminin biri pozitif, diğeri negatif iki gerçel kökü varsa m nin alabileceği değerler kümesi aşağıdakilerden hangisidir?

A) $(1, \infty)$ B) $(-2, 2)$

C) $(-1, 0) \cup (1, \infty)$ D) $(-2, 1) \cup (2, \infty)$

E) $(-2, 0) \cup (1, \infty)$

MT2.8



$m(\widehat{AOP}) = \theta$

$m(\widehat{AOP'}) = -\theta$

Şekildeki O merkezli birim çember üzerindeki P ve P' noktaları Ox eksenine göre birbirinin simetriğidir.

Buna göre, P' noktası aşağıdakilerden hangisiyle <u>ifade edilemez</u>?

A) $(\cos(-\theta), \sin(-\theta))$

B) $(\cos(-\theta), \sin\theta)$

C) $(\cos\theta, -\sin\theta)$

D) $(\cos\theta, \sin(2\pi-\theta))$

E) $(\cos(2\pi-\theta), -\sin\theta)$

MT2.9

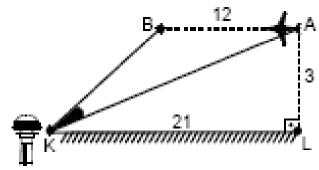
$$\frac{\sin 2a}{1-\cos 2a}$$

ifadesinin sadeleştirilmiş biçimi aşağıdakilerden hangisidir?

A) $\sin a$ B) $\cos a$ C) $\tan a$

D) $\cot a$ E) $\sin a + \cos a$

MT2.10



$AL \perp KL$

$BA \,//\, KL$

$|AL| = 3$ km

$|BA| = 12$ km

$|KL| = 21$ km

K noktasındaki kontrol kulesinde bulunan bir görevli, yerden 3 km yükseklikte yere paralel uçan bir uçağın, A noktasından B noktasına kadar 12 km lik hareketini radarla izliyor.

A noktasının yerdeki dik izdüşümü L noktası ve $|KL| = 21$ km olduğuna göre, radarın taradığı AKB açısının tanjantı kaçtır?

A) $\frac{3}{7}$ B) $\frac{4}{9}$ C) $\frac{2}{11}$

D) $\frac{3}{13}$ E) $\frac{7}{17}$

MT2.11

$f : \left(\frac{-1}{3}, \infty\right) \to R$ fonksiyonu

$f(x) = \log_3(3x+1)$

ile tanımlanıyor.

Buna göre, ters fonksiyonu belirten $f^{-1}(x)$ aşağıdakilerden hangisidir?

A) $f^{-1}(x) = 3^x$ B) $f^{-1}(x) = 3^x + 1$

C) $f^{-1}(x) = \log(3x+1)$ D) $f^{-1}(x) = \frac{3^x - 1}{3}$

E) $f^{-1}(x) = \frac{x^3+1}{3}$

MT2.12

$$f(x) = \begin{cases} \frac{|x|}{x}, & x = 0 \text{ ise} \\ 3, & x = 0 \text{ ise} \end{cases}$$

fonksiyonu için,

$\lim\limits_{x \to 0^+} f(x) = a$

$\lim\limits_{x \to 0^-} f(x) = b$

olduğuna göre, $a - b$ kaçtır?

A) -2 B) -1 C) 0 D) 1 E) 2

MT2.13

$$s_n = \sum_{k=1}^{n} \frac{k}{n^2}$$

olduğuna göre, $\lim\limits_{n \to \infty} s_n$ kaçtır?

A) $\frac{1}{2}$ B) $\frac{2}{3}$ C) 0 D) 1 E) 2

MT2.14

$f : R \to R$ her noktada türevli bir fonksiyon ve $f'(1) = 3$ olduğuna göre,

$$\lim_{h \to 0} \frac{f(1+2h) - f(1-3h)}{h}$$

kaçtır?

A) 15 B) 12 C) 9 D) 6 E) 3

MT2.15

$P(x)$ polinom fonksiyonunun türevi $P'(x)$ ve

$$P(x) - P'(x) = 2x^2 + 3x - 1$$

olduğuna göre, $P(x)$ in katsayılarının toplamı kaçtır?

A) 11 B) 12 C) 13 D) 14 E) 15

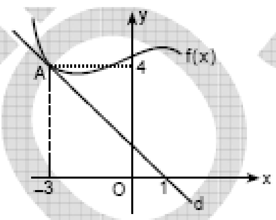
MT2.16

$$f(x) = \frac{2x^3}{3} - \frac{x^2}{2} + 5$$

fonksiyonu aşağıdaki aralıkların hangisinde <u>azalandır</u>?

A) $\left(\frac{-3}{2}, -1\right)$ B) $\left(-1, \frac{-1}{2}\right)$ C) $\left(\frac{-1}{2}, 0\right)$ D) $\left(0, \frac{1}{2}\right)$ E) $\left(\frac{1}{2}, \frac{3}{2}\right)$

MT2.17



Şekildeki d doğrusu, $f(x)$ fonksiyonunun grafiğine A noktasında teğettir.

$h(x) = x \cdot f(x)$ olduğuna göre, $h'(-3)$ kaçtır?

A) −4 B) −2 C) 0 D) 2 E) 7

MT2.18

$$\int_{\frac{\pi}{2}}^{\pi} (\sin x + \cos x)\,dx$$

integralinde $t = \pi - x$ dönüşümü yapılırsa aşağıdaki integrallerden hangisi elde edilir?

A) $\int_0^{\frac{\pi}{2}} (\sin t + \cos t)\,dt$ B) $\int_0^{\frac{\pi}{2}} (\sin t - \cos t)\,dt$

C) $\int_{\frac{\pi}{2}}^{\pi} (\sin t - \cos t)\,dt$ D) $\int_{\frac{\pi}{2}}^{\pi} (\cos t - \sin t)\,dt$

E) $\int_{\frac{-\pi}{2}}^{0} (\sin t - \cos t)\,dt$

MT2.19

$f : R \to R$ fonksiyonu her noktada türevli ve

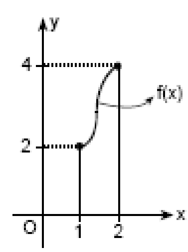
$$f'(x) = x + 1$$

$$f(2) = -1$$

olduğuna göre, $f(0)$ kaçtır?

A) −5 B) −4 C) −2 D) −1 E) 0

MT2.20



Şekilde grafiği verilen bire bir ve örten $f : [1,2] \to [2,4]$ fonksiyonunun tersi f^{-1} dir.

Buna göre,

$$\int_1^2 f(x)\,dx + \int_2^4 f^{-1}(x)\,dx$$

toplamı kaçtır?

A) 2 B) 4 C) 6 D) 8 E) 10

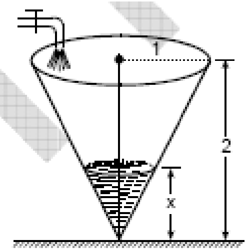
MT2.21

$$\begin{vmatrix} \log_2 8 & \log_4 5 \\ \log_5 4 & \dfrac{1}{\log_{27} 3} \end{vmatrix}$$

determinantının değeri kaçtır?

A) 10 B) 9 C) 8 D) 6 E) 5

MT2.22



Şekildeki gibi, taban yarıçapı 1 metre, yüksekliği 2 metre olan dik koni biçimindeki bir su deposuna bir musluktan sabit hızla su akıtılıyor.

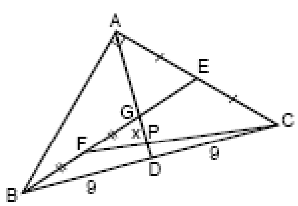
Depoda biriken suyun derinliği x metre olduğunda, depoda biriken suyun hacmi x türünden kaç metreküp olur?

A) $\frac{\pi x^3}{12}$ B) $\frac{\pi x^3}{9}$ C) $\frac{\pi x^3}{6}$

D) $\frac{\pi x^3}{4}$ E) $\frac{\pi x^3}{3}$

MT2.23

22.



ABC bir dik üçgen

$m(\widehat{BAC}) = 90°$

$|AE| = |EC|$

$|BD| = |DC| = 9$ cm

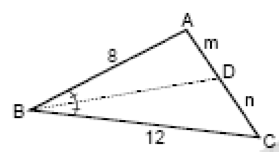
$|BF| = |FG|$

$|GP| = x$

Yukarıdaki verilere göre, x kaç cm dir?

A) 1 B) 2 C) 3 D) $\frac{3}{2}$ E) $\frac{5}{2}$

MT2.24



ABC bir üçgen

$[BD]$ açıortay

$|AB| = 8$ cm

$|BC| = 12$ cm

$|AD| = m$ cm

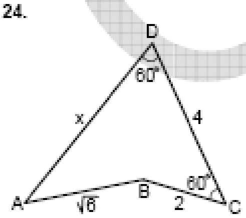
$|DC| = n$ cm

Yukarıdaki şekilde m ve n birer tamsayı olduğuna göre, ABC üçgeninin çevre uzunluğu en çok kaç cm olabilir?

A) 28 B) 32 C) 35 D) 38 E) 40

MT2.25

24.



$m(\widehat{ADC}) = m(\widehat{BCD}) = 60°$

$|AB| = \sqrt{6}$ cm

$|BC| = 2$ cm

$|CD| = 4$ cm

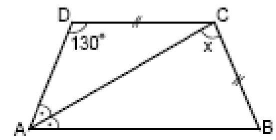
$|AD| = x$

Yukarıdaki verilere göre, x kaç cm dir?

A) 5 B) 6 C) $6-\sqrt{3}$

D) $2+\sqrt{6}$ E) $3+\sqrt{3}$

MT2.26



AB // DC

[AC] açıortay

$|DC| = |BC|$

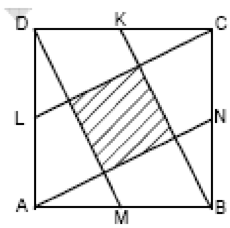
$m(\widehat{ADC}) = 130°$

$m(\widehat{ACB}) = x$

Yukarıdaki verilere göre, x kaç derecedir?

A) 105 B) 115 C) 125 D) 130 E) 135

MT2.27



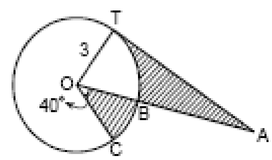
Şekildeki ABCD karesinin kenarları üzerindeki K, L, M, N noktalarının her biri, üzerinde bulunduğu kenarın orta noktasıdır.

A(ABCD) = 4 br² olduğuna göre, taralı alan kaç br² dir?

A) $\frac{1}{2}$ B) $\frac{1}{4}$ C) $\frac{4}{5}$

D) $\frac{2}{5}$ E) $\frac{1}{5}$

MT2.28



$m(\widehat{COB}) = 40°$

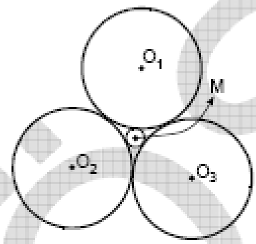
$|OT| = 3$ cm

Şekildeki AT doğrusu O merkezli çembere T noktasında teğettir ve $|AT|$ uzunluğu TBC yayının uzunluğuna eşittir.

Buna göre, taralı alanların toplamı kaç cm² dir?

A) 8π B) 6π C) 5π D) 4π E) 2π

MT2.29



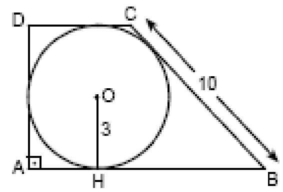
O_1, O_2, O_3 ve M merkezli çemberler birbirlerine şekildeki gibi teğettir.

O_1, O_2 ve O_3 merkezli çemberlerin yarıçapları r cm, M merkezli çemberin yarıçapı da 1 cm olduğuna göre, r kaçtır?

A) $\sqrt{3}$ B) $1+\sqrt{3}$ C) $2+2\sqrt{3}$

D) $3+2\sqrt{3}$ E) $3+3\sqrt{3}$

MT2.30



ABCD, O merkezli çemberin teğetler dörtgeni

AB // DC

DA ⊥ AB

$|BC| = 10$ cm

$|OH| = 3$ cm

Yukarıdaki verilere göre, ABCD teğetler dörtgeninin alanı kaç cm^2 dir?

A) 50 B) 48 C) 46 D) 44 E) 42

APPENDIX C

2007 SST MATHEMATICS SECTION I ITEMS

MT1.1 $1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}$ **işleminin sonucu kaçtır?** A) $\frac{5}{8}$ B) $\frac{3}{8}$ C) $\frac{1}{8}$ D) $\frac{3}{4}$ E) $\frac{1}{4}$	MT1.2 $\dfrac{\left(1-\frac{1}{2}\right)\left(1-\frac{1}{4}\right)}{\left(1+\frac{1}{2}\right)\left(1+\frac{1}{4}\right)}$ **işleminin sonucu kaçtır?** A) $\frac{1}{7}$ B) $\frac{1}{5}$ C) $\frac{1}{4}$ D) $\frac{3}{4}$ E) $\frac{2}{3}$
MT1.3 $\frac{1}{20}$ **kesrinin ondalık gösterimi aşağıdakilerden hangisidir?** A) 0,01 B) 0,02 C) 0,05 D) 0,2 E) 1,2	MT1.4 $\dfrac{(-1)^2+5-2^2}{35-21}$ **işleminin sonucu kaçtır?** A) $\frac{-3}{14}$ B) $\frac{-1}{14}$ C) $\frac{5}{14}$ D) $\frac{1}{7}$ E) $\frac{4}{7}$
MT1.5 $a=0,\overline{2}=0,2222\cdots2\cdots$ **devirli ondalık açılımıyla verilen a sayısı için $\sqrt{a}$ aşağıdakilerden hangisidir?** A) $\frac{\sqrt{5}}{2}$ B) $\frac{\sqrt{3}}{2}$ C) $\frac{\sqrt{2}}{2}$ D) $\frac{\sqrt{5}}{3}$ E) $\frac{\sqrt{2}}{3}$	MT1.6 $\left(\sqrt{2}-\sqrt{5}\right)^2+2\sqrt{10}+3$ **işleminin sonucu kaçtır?** A) $\sqrt{10}$ B) $2\sqrt{5}$ C) $5\sqrt{2}$ D) 10 E) 13

MT1.7

$$3^{\frac{-1}{2}} \cdot \sqrt{27}$$

işleminin sonucu kaçtır?

A) 3 B) 9 C) $\sqrt{3}$ D) $3\sqrt{3}$ E) $\frac{\sqrt{3}}{3}$

MT1.8

$$\frac{3^{2x} - 2 \cdot 3^{x+y} + 3^{2y}}{3^{2x} - 3^{x+y}}$$

işleminin sonucu aşağıdakilerden hangisidir?

A) $3^x - 3^y$ B) $3^x + 3^y$ C) $1 + 3^{y-x}$

D) $1 - 3^{x+y}$ E) $1 - 3^{y-x}$

MT1.9

1 den farklı a ve b pozitif gerçel sayıları için

$$ab = a^b$$

$$\frac{a}{b} = a^{2b}$$

olduğuna göre, b kaçtır?

A) $\frac{2}{3}$ B) $\frac{3}{4}$ C) $\frac{4}{5}$

D) $\frac{5}{6}$ E) $\frac{6}{7}$

MT1.10

a ve b pozitif tam sayılar olmak üzere,

$$a^2 - 2ab - 3b^2 = 0$$

olduğuna göre, $a + b$ toplamının en küçük değeri kaçtır?

A) 7 B) 6 C) 5 D) 4 E) 3

MT1.11

$\{1, 2, 3, 4, 5\}$ kümesinin birbirinden farklı a, b ve c elemanları için

$$3a - b - 2c$$

ifadesinin en büyük değeri kaçtır?

A) 10 B) 11 C) 12 D) 14 E) 15

MT1.12

$$(x^2 - x - 2)(x + 5) = 0$$

denkleminin köklerinin toplamı kaçtır?

A) 3 B) 1 C) -2 D) -4 E) -6

MT1.13

m ve n pozitif tam sayılarının ortak bölenlerinin en büyüğü $\text{OBEB}(m, n) = 6$ ve ortak katlarının en küçüğü $\text{OKEK}(m, n) = 60$ tır.

$m + n = 42$ olduğuna göre, $|m - n|$ kaçtır?

A) 26 B) 24 C) 22 D) 20 E) 18

MT1.14

n doğal sayı olmak üzere, $2^{2^n} + 1$ biçiminde yazılabilen asal sayılara **Fermat asal sayıları** denir.

Buna göre, aşağıdakilerden hangisi Fermat asal sayısıdır?

A) 7 B) 11 C) 13 D) 17 E) 23

MT1.15

$(x-2)(x+2)(x+5)=(x-1)(x+1)(x+4)$

denklemiyle aşağıdaki denklemlerden hangisinin çözüm kümesi aynıdır?

A) $x^3+5x^2+4x=0$

B) $x^2-3x-16=0$

C) $x^2-4x+24=0$

D) $3x+16=0$

E) $5x-4=0$

MT1.16

Badem, çekirdek, fıstık ve leblebi karıştırılarak bir kuruyemiş paketi hazırlanmıştır. Aşağıdaki tabloda bu paketteki çekirdek, fıstık ve leblebinin ağırlıklarıyla çekirdeğin ağırlıkça yüzde oranı verilmiştir.

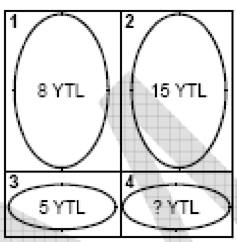
	Ağırlığı (g)	Yüzde oranı (%)
Badem		
Çekirdek	500	40
Fıstık	300	
Leblebi	250	

Bu paketteki bademin ağırlıkça yüzde oranı kaçtır?

A) 12 B) 15 C) 16 D) 18 E) 24

MT1.17

Dört gözlü bir yazar kasa çekmecesinin 1 ve 2 numaralı gözlerindeki paraların tutarı birbirine eşittir; 3 ve 4 numaralı gözlerindeki paraların tutarı da birbirine eşittir. Bu çekmecenin 1 ve 3 numaralı gözlerinin her birine a YTL, 2 ve 4 numaralı gözlerinin her birine de b YTL tutarında para konulunca şekilde belirtilen tutarlar elde ediliyor.



Buna göre, 4 numaralı gözde son durumda kaç YTL vardır?

A) 7 B) 10 C) 12 D) 13 E) 14

MT1.18

Bir tüccar, tanesi 45 YTL den belirli sayıda gömlek satın alıyor. Kendisine verilen faturada, ödenen miktarın ilk ve son rakamları silik çıktığı için bu tutarın yalnızca •92• biçiminde dört basamaklı bir sayı olduğu okunabiliyor.

Tüccarın <u>tek sayıda</u> gömlek aldığı bilindiğine göre, silik çıkan iki rakamın toplamı kaçtır?

A) 6 B) 7 C) 8 D) 9 E) 10

MT1.19

Bir aracın duruş mesafesi, frene basıldığı andaki hızının karesiyle doğru orantılıdır.

Bu araç saatte 60 km hızla giderken duruş mesafesi 20 m olduğuna göre, saatte 90 km hızla giderken duruş mesafesi kaç m dir?

A) 30 B) 45 C) 50 D) 60 E) 72

MT1.20

Dört kardeş 114 YTL yi paylaşıyor. Bu paylaşmada birinci kardeş ikinciden 1 YTL, ikinci üçüncüden 2 YTL, üçüncü dördüncüden 3 YTL fazla alıyor.

Buna göre, <u>en fazla</u> para alan kaç YTL almıştır?

A) 27 B) 28 C) 29 D) 31 E) 38

MT1.21

Sabit bir hızla yürüyen İrem, evden okula giderken yolun $\frac{1}{3}$ ünü yürüdüğünde matematik defterini yanına almadığını fark ediyor.

İrem yoluna devam ederse dersin başlamasından 4 dakika önce, eve dönerek defterini alıp tekrar yola çıkarsa dersin başlamasından 4 dakika sonra okula varacağına göre, ev ile okul arasını kaç dakikada almaktadır?

(Dönüşlerdeki zaman kayıpları önemsenmeyecektir.)

A) 10 B) 12 C) 14 D) 15 E) 16

MT1.22

Bir müşteri aldığı tişört için kasiyere bir miktar para vermiştir. Kasiyer, tişört fiyatındaki YTL ve YKr bölümlerini karıştırmış (örneğin tişört 16,05 YTL ise kasiyer, fiyatı 5,16 YTL olarak görmüş) ve müşteriye 4,80 YTL yerine yanlışlıkla 19,65 YTL para üstü vermiştir.

Tişörtün gerçek fiyatıyla kasiyerin gördüğü fiyatın toplamı 55,55 YTL olduğuna göre, müşteri kasiyere kaç YTL vermiştir?

A) 60 B) 55 C) 50 D) 45 E) 40

MT1.23

$A = \{-2, -1, 0, 1\}$

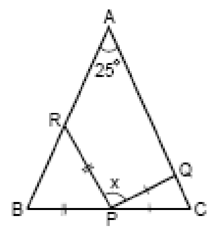
$B = \{-1, 0, 1, 2, 3, 4\}$

kümeleri veriliyor.

$A \times B$ kartezyen çarpımından alınan bir elemanın (a, a) biçiminde olma olasılığı kaçtır?

A) $\frac{1}{4}$ B) $\frac{1}{6}$ C) $\frac{1}{8}$ D) $\frac{1}{12}$ E) $\frac{5}{24}$

MT1.24



ABC bir üçgen

$|BP| = |PR|$

$|CP| = |PQ|$

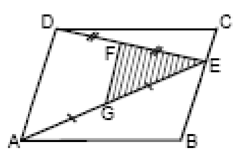
$m(\widehat{BAC}) = 25°$

$m(\widehat{RPQ}) = x$

Yukarıdaki verilere göre, x kaç derecedir?

A) 150 B) 135 C) 130 D) 120 E) 108

MT1.25



ABCD bir paralelkenar

$|DF| = |FE|$

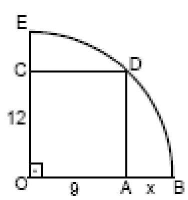
$|AG| = |GE|$

Şekildeki ABCD paralelkenarının alanı $72\ cm^2$ dir.

Buna göre, taralı EFG üçgeninin alanı kaç cm^2 dir?

A) 9 B) 10 C) 12 D) 16 E) 18

MT1.26



OADC bir dikdörtgen

$|OC| = 12$ cm

$|OA| = 9$ cm

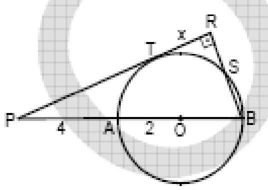
$|AB| = x$

Şekildeki E, D ve B noktaları O merkezli çeyrek çemberin üzerindedir.

Buna göre, x kaç cm dir?

A) 10 B) 9 C) 8 D) 7 E) 6

MT1.27



PR doğrusu O merkezli çembere T noktasında teğet

$PR \perp RB$

$|PA| = 4$ cm

$|AO| = 2$ cm

$|TR| = x$

Yukarıdaki verilere göre, x kaç cm dir?

A) $\frac{4}{3}\sqrt{2}$ B) $\frac{5}{4}\sqrt{2}$ C) $\frac{3}{2}\sqrt{3}$

D) $\frac{5}{3}\sqrt{3}$ E) $\frac{2}{3}\sqrt{5}$

MT1.28

XOY dik koordinat sistemiyle verilen düzlemde $A_0(-1, 2)$ noktasından başlayıp her seferinde x koordinatını 1 birim, y koordinatını 2 birim artırarak $A_1, A_2, \dots, A_n$ noktaları işaretleniyor.

A_n noktası $y = 3x$ doğrusu üzerinde olduğuna göre, n kaçtır?

A) 8 B) 7 C) 6 D) 5 E) 4

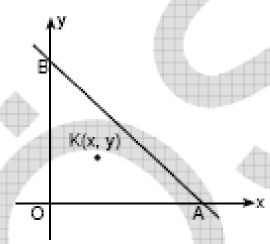
MT1.29

Dik koordinat düzleminde denklemi $x + y = 3$ olan doğrunun, Oy eksenine göre simetriğinin denklemi aşağıdakilerden hangisidir?

A) $-x + y = 3$ B) $x - y = 3$ C) $-x - y = 3$

D) $x + 2y = 1$ E) $2x + y = 1$

MT1.30



Şekildeki dik koordinat sisteminin eksenleri üzerindeki A ve B noktalarını birleştiren $[AB]$ doğru parçasının uzunluğu 12 cm dir.

OAB üçgeninin kenarortayları $K(x, y)$ noktasında kesiştiğine göre, $x^2 + y^2$ toplamı kaçtır?

A) 12 B) 15 C) 16 D) 18 E) 25

APPENDIX D

2007 SST MATHEMATICS SECTION II ITEMS

MT2.1 $\dfrac{\dfrac{1}{x}-x}{x+x^2}\cdot\dfrac{x^2}{1-2x+x^2}$ **ifadesinin sadeleştirilmiş biçimi aşağıdakilerden hangisidir?** A) $\dfrac{1}{x^2}$ B) $\dfrac{x}{1-x}$ C) $\dfrac{1}{1-x}$ D) $\dfrac{1}{1+x}$ E) $\dfrac{1-x}{1+x}$	MT2.2 $\dfrac{x^2+x+1}{2x^2+5x}:\dfrac{x^3-1}{2x^2+3x-5}$ **ifadesinin sadeleştirilmiş biçimi aşağıdakilerden hangisidir?** A) $\dfrac{1}{x}$ B) $\dfrac{1}{2-x}$ C) $\dfrac{2}{1+x}$ D) x E) $x+1$
MT2.3 Karmaşık sayılar kümesi üzerinde $*$ işlemi, $z_1 * z_2 = z_1 + z_2 + \lvert z_1 z_2 \rvert$ biçiminde tanımlanıyor. **Buna göre, $(1-2i)*(2+i)$ işleminin sonucu nedir?** A) $1+8i$ B) $1-8i$ C) $8+i$ D) $8-i$ E) $2-i$	MT2.4 $\dfrac{\sin 10^\circ \cos 40^\circ + \cos 10^\circ \sin 40^\circ}{\cos 50^\circ \cos 10^\circ + \sin 50^\circ \sin 10^\circ}$ **işleminin sonucu kaçtır?** A) $\sqrt{2}$ B) $\sqrt{3}$ C) $\dfrac{\sqrt{3}}{2}$ D) $\dfrac{1}{2}$ E) 1
MT2.5 $\dfrac{\cos 2a}{1-\tan^2 a}$ **ifadesinin sadeleştirilmiş biçimi aşağıdakilerden hangisidir?** A) $\sin^2 a$ B) $\cos^2 a$ C) $\cot^2 a$ D) $1+\sin^2 a$ E) $1+\tan^2 a$	MT2.6 $\left(\sin\dfrac{\pi}{12}+\cos\dfrac{\pi}{12}\right)^2$ **ifadesinin değeri kaçtır?** A) $\dfrac{1}{2}$ B) $\dfrac{3}{2}$ C) $\dfrac{5}{2}$ D) $-1+\sqrt{3}$ E) $1+\sqrt{3}$

MT2.7

$\log_2\left(\log_3(5x+6)\right)=2$

olduğuna göre, x kaçtır?

A) 6 B) 8 C) 9 D) 15 E) 18

MT2.8

$n \geq 1$ için

$$a_n=\sum_{k=1}^{n}\frac{1}{k(k+1)}$$

olduğuna göre, a_{99} aşağıdakilerden hangisidir?

A) $\frac{50}{49}$ B) $\frac{49}{50}$ C) $\frac{98}{99}$

D) $\frac{100}{99}$ E) $\frac{99}{100}$

MT2.9

$f(x)=\left||x-3|-2\right|$ fonksiyonunun grafiğiyle $g(x)=4$ fonksiyonunun grafiğinin kesim noktalarının apsislerinin toplamı kaçtır?

A) 16 B) 14 C) 10 D) 8 E) 6

MT2.10

$f(x)=2\sqrt{1-x^2}$

ile verilen f fonksiyonunun gerçel sayılardaki en geniş tanım kümesi T ve görüntü kümesi $G=\{f(x)\,|\,x\in T\}$ olduğuna göre, $T\cap G$ kesişim kümesi aşağıdaki aralıklardan hangisine eşittir?

A) $[0,1]$ B) $[1,2]$ C) $[2,3]$

D) $[0,\sqrt{2}]$ E) $[1,\sqrt{2}]$

MT2.11

R den R ye $f(x)=3^{x+2}$ ile tanımlı f fonksiyonu için, $f(a+b-1)$ ifadesi aşağıdakilerden hangisine eşittir?

A) $\frac{f(a+b)}{9}$ B) $\frac{f(a+b)}{27}$

C) $\frac{f(a)\cdot f(b)}{9}$ D) $\frac{f(a)\cdot f(b)}{27}$

E) $\frac{f(a)\cdot f(b)}{81}$

MT2.12

R den R ye

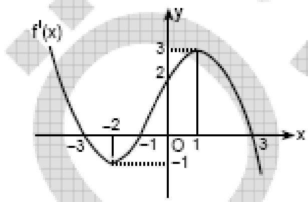
$$f(x)=\begin{cases} x^2 & , \quad x<3 \text{ ise} \\ 3 & , \quad x=3 \text{ ise} \\ x+a & , \quad x>3 \text{ ise} \end{cases}$$

ile tanımlanan f fonksiyonunun $x=3$ noktasında limitinin olması için a kaç olmalıdır?

A) 4 B) 6 C) 7 D) 8 E) 9

MT2.13

Aşağıda, her noktada türevlenebilir bir f fonksiyonunun <u>türevinin</u> (f' nün) grafiği verilmiştir.



Yukarıdaki verilere uygun olarak alınacak her <u>f fonksiyonu</u> için aşağıdakilerden hangisi kesinlikle doğrudur?

A) $-2 < x < -1$ aralığında artandır.

B) $0 < x < 3$ aralığında azalandır.

C) $x = 1$ de bir yerel maksimumu vardır.

D) $x = -1$ de bir yerel maksimumu vardır.

E) $x = -3$ te bir yerel maksimumu vardır.

MT2.14

$$\lim_{x \to 0^+} \frac{1-\cos\sqrt{x}}{x}$$

limitinin değeri kaçtır?

A) 0 B) $\frac{1}{2}$ C) 1 D) 2 E) $\sqrt{2}$

MT2.15

Gerçel sayılar kümesi üzerinde, tanımlı ve türevlenebilir bir f fonksiyonu için

$$f(x+y) = f(x) + f(y) + xy$$

$$\lim_{h \to 0} \frac{f(h)}{h} = 3$$

olduğuna göre, $f'(1)$ kaçtır?

A) 2 B) 3 C) 4 D) 5 E) 6

MT2.16

Gerçel sayılar kümesi üzerinde, tanımlı ve türevlenebilir bir f fonksiyonu için $f(0) = f'(0) = 4$ olduğuna göre,

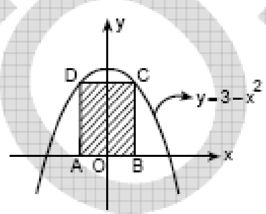
$$g(x) = f(x \cdot f(x))$$

ile tanımlanan g fonksiyonu için $g'(0)$ kaçtır?

A) 0 B) 4 C) 8 D) 12 E) 16

MT2.17

A ve B noktaları Ox ekseni üzerinde, C ve D noktaları ise $y = 3 - x^2$ parabolü üzerinde pozitif ordinatlı noktalar olmak üzere şekildeki gibi ABCD dikdörtgenleri oluşturuluyor.



Bu dikdörtgenlerden alanı <u>en büyük</u> olanın alanı kaç birim karedir?

A) 2 B) 3 C) 4 D) 5 E) 6

MT2.18

$$\int_0^1 3x\sqrt{3+x^2}\,dx$$

integralinin değeri kaçtır?

A) $1+\sqrt{3}$ B) $2-2\sqrt{3}$ C) $2+\sqrt{3}$

D) $4-\sqrt{3}$ E) $8-3\sqrt{3}$

MT2.19

$$\int_0^1 \frac{x^2}{x+1}\,dx$$

integralinin değeri kaçtır?

A) $\frac{-1}{2}+\ln 2$ B) $-1+\ln 2$ C) $\ln 2$

D) $2\ln 2$ E) $1+2\ln 2$

MT2.20

$x^2 = 2y$

$y^2 = 2x$

eğrileriyle sınırlanan bölgenin alanı kaç birim karedir?

A) $\frac{5}{2}$ B) $\frac{1}{3}$ C) $\frac{2}{3}$

D) $\frac{4}{3}$ E) $\frac{5}{4}$

MT2.21

$A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$

$B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

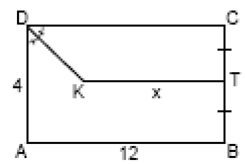
matrisleri için $A \cdot X = B$ denklemini sağlayan X matrisi aşağıdakilerden hangisidir?

A) $\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$ B) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

C) $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ D) $\begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix}$

E) $\begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix}$

MT2.22



ABCD bir dikdörtgen

$KT \parallel AB$

$m(\widehat{ADK}) = m(\widehat{KDC})$

$|CT| = |TB|$

$|AD| = 4$ cm

$|AB| = 12$ cm

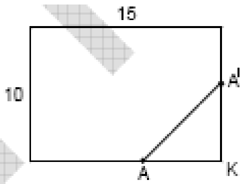
$|KT| = x$

Yukarıdaki verilere göre, x kaç cm dir?

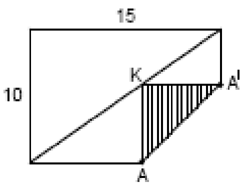
A) 8,5 B) 9 C) 9,5 D) 10 E) 10,5

MT2.23

Şekil I



Şekil II

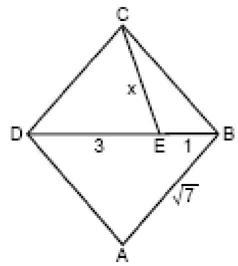


Boyutları 15 cm ve 10 cm olan Şekil I deki dikdörtgen biçiminde bir karton, K köşesine eşit uzaklıkta olan A ve A' noktalarını birleştiren AA' doğrusu boyunca Şekil II deki gibi katlandığında K köşesi dikdörtgenin köşegeni üzerine geliyor.

Katlanan AA'K üçgensel bölgesinin alanı kaç cm^2 dir?

A) 18 B) 20 C) 25 D) 30 E) 32

MT2.24



ABCD bir eşkenar dörtgen

$|AB| = \sqrt{7}$ cm

$|DE| = 3$ cm

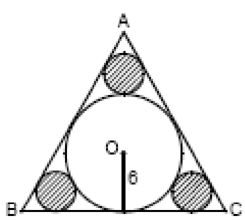
$|EB| = 1$ cm

$|CE| = x$

Yukarıdaki verilere göre, x kaç cm dir?

A) 1 B) 2 C) $\sqrt{2}$ D) $\sqrt{3}$ E) $\sqrt{5}$

MT2.25



Şekildeki ABC üçgeni eşkenar üçgendir ve O merkezli çember ABC üçgeninin iç teğet çemberidir. Küçük çemberler de bu çembere ve üçgenin kenarlarına teğettir.

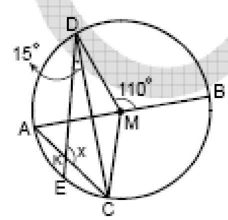
O merkezli çemberin yarıçapı 6 cm olduğuna göre küçük çemberlerin alanları toplamı kaç cm^2 dir?

A) 6π B) 9π C) 12π

D) 15π E) 18π

MT2.26

M merkezli bir çemberin [AB] çapının ayırdığı farklı yaylar üzerinde C ve D noktaları alınıyor. [AC] kirişi üzerinde alınan bir K noktası için DK doğrusu, çemberi E noktasında kesiyor.



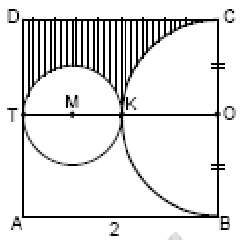
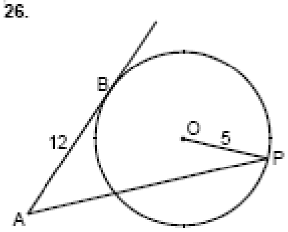
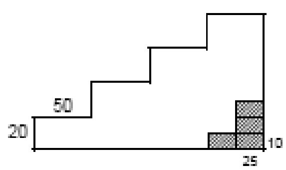
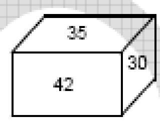
$m(\widehat{EDC}) = 15°$

$m(\widehat{DMB}) = 110°$

$m(\widehat{DKC}) = x$

Yukarıdaki verilere göre, x kaç derecedir?

A) 130 B) 125 C) 120 D) 115 E) 105

MT2.27	MT2.28
 ABCD bir kare $\|OB\| = \|OC\|$ TO // AB $\|AB\| = 2$ cm Şekildeki M merkezli çember $[AD]$ kenarına T noktasında ve O merkezli, $[BC]$ çaplı yarı çembere K noktasında teğettir. **Buna göre, taralı bölgenin alanı kaç cm^2 dir?** A) $2-\frac{3\pi}{8}$ B) $2-\frac{5\pi}{8}$ C) $2-\frac{3\pi}{7}$ D) $4-\frac{3\pi}{8}$ E) $4-\frac{5\pi}{7}$	**26.**  AB doğrusu O merkezli çembere B noktasında teğet $\|OP\| = 5$ cm $\|AB\| = 12$ cm Şekildeki P noktası çember üzerinde değişmektedir. **Buna göre $\|AP\|$ uzunluğunun <u>en büyük</u> değeri kaç cm dir?** A) 22 B) 20 C) 19 D) 18 E) 17
MT2.29	MT2.30
Basamak yüksekliği 20 cm, basamak genişliği 50 cm olan aşağıdaki merdivenin yan yüzü, boyutları 25 cm ve 10 cm olan dikdörtgen biçimindeki fayanslarla kaplanacaktır.  **Bu iş için kaç tane fayans kullanılır?** A) 40 B) 38 C) 36 D) 32 E) 28	 Şekildeki dikdörtgenler prizmasının üç farklı yüzünün alanları cm^2 türünden üzerlerine yazılmıştır. **Bu prizmanın hacmi kaç cm^3 tür?** A) 200 B) 210 C) 240 D) 260 E) 280

APPENDIX E

EXAMPLES FOR A RELEASED PISA 2003 PROBLEM SOLVING ITEM

EXAMPLE 1.

ENERGY NEEDS

This unit asks two questions of students. The first problem, shown below, is about selecting suitable food to meet the energy needs of a person in Zedland. Success indicated that a student was able to look up needed information for solving a problem. This item's demands were below those associated with Level 1.

Daily energy needs recommended for adults

		Men	Women
Age (years)	**Activity level**	**Energy needed** *(kJ)*	**Energy needed** *(kJ)*
From 18 to 29	Light Moderate Heavy	10660 11080 14420	8360 8780 9820
From 30 to 59	Light Moderate Heavy	10450 12120 14210	8570 8990 9790
60 and above	Light Moderate Heavy	8780 10240 11910	7500 7940 8780

Activity level according to occupation

Light:	*Moderate:*	*Heavy:*
Indoor sales person Office worker Housewife	Teacher Outdoor salesperson Nurse	Construction worker Labourer Sportsperson

Jane Gibbs is a 19-year old high jumper. One evening, some of Jane's friends invite her out for dinner at a restaurant. Here is the menu:

	MENU	**Jane's estimate of energy per serving** *(kJ)*
Soups:	Tomato Soup	355
	Cream of Mushroom Soup	585
Main courses:	Mexican Chicken	960
	Caribbean Ginger Chicken	795
	Pork and Sage Kebabs	920
Salads:	Potato salad	750
	Spinach, Apricot and Hazelnut Salad	335
	Couscous Salad	480
Desserts:	Apple and Raspberry Crumble	1380
	Ginger Cheesecake	1005
	Carrot Cake	565
Milkshakes:	Chocolate	1590
	Vanilla	1470

The restaurant also has a special fixed price menu.

Fixed Price Menu
50 zeds
Tomato Soup
Caribbean Ginger Chicken
Carrot Cake

ENERGY NEEDS – Question 2

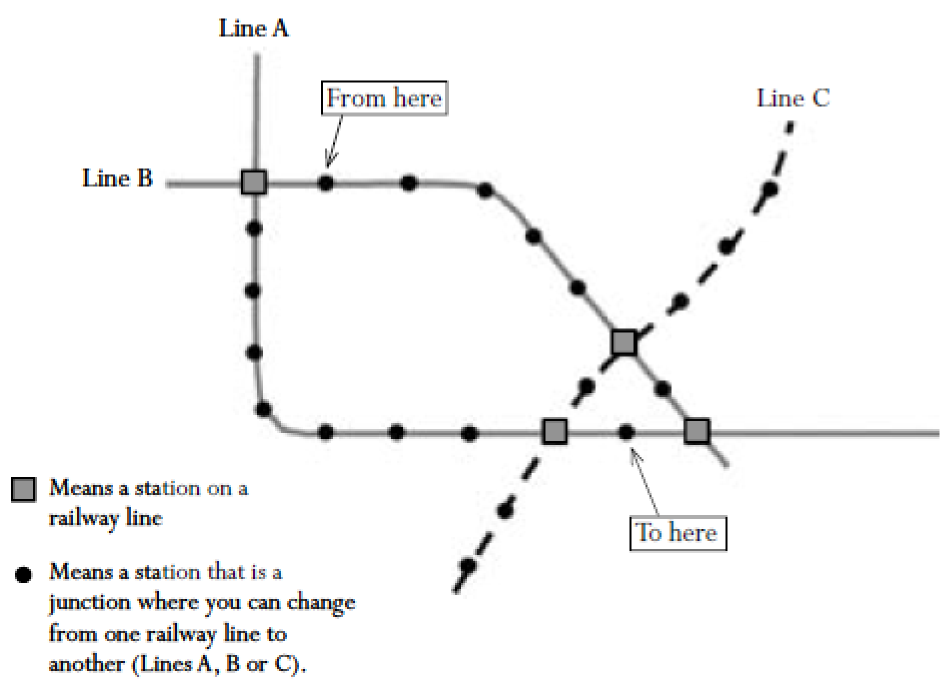
Jane keeps a record of what she eats each day. Before dinner on that day her total intake of energy had been 7520 kJ.

Jane does **not** want her total energy intake to go **below or above her recommended daily amount** by more than 500 kJ.

Decide whether the special "Fixed Price Menu" will allow Jane to stay within ±500 kJ of her recommended energy needs. Show your work.

EXAMPLE 2.

The following diagram shows part of the transport system of a city in Zedland, with three railway lines. It shows where you are at present, and where you have to go.



The fare is based on the number of stations travelled (not counting the station where you start your journey). Each station travelled costs 1 zed.

The time taken to travel between two adjacent stations is about 2 minutes.

The time taken to change from one railway line to another at a junction is about 5 minutes.

TRANSIT SYSTEM – Question 1

The diagram indicates a station where you are currently at ("From here"), and the station where you want to go ("To here"). **Mark on the diagram** the best route in terms of cost and time, and indicate below the fare you have to pay, and the approximate time for the journey.

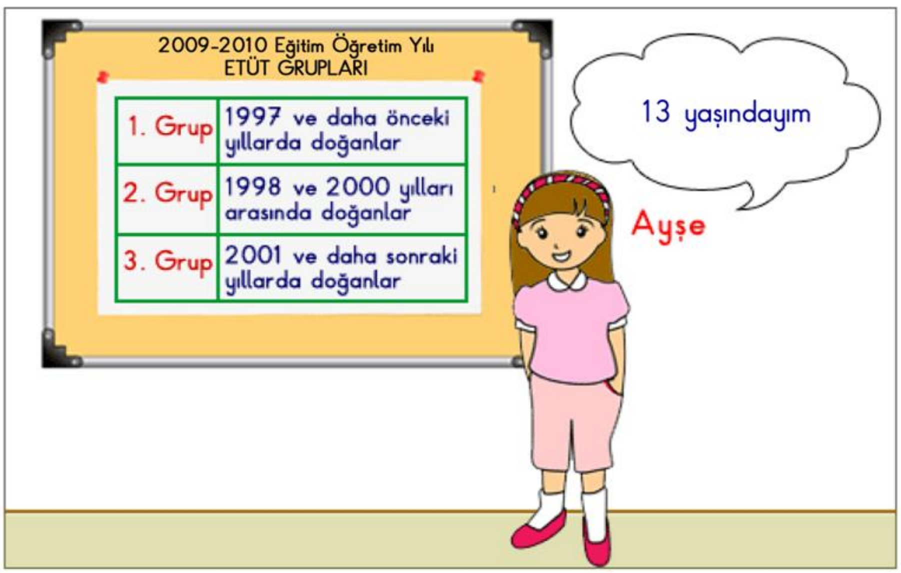
Fare: .... zeds.

Approximate time for journey: .... minutes.

APPENDIX F

EXAMPLE FOR A RELEASED CITO TURKEY PMS PROBLEM SOLVING ITEM

Ayşe 13 yaşındadır. Hangi etüt grubunda yer alması gerekir? Neden?



APPENDIX G

P50 AND P80 VALUES OF ITEMS

Item	Disc	P50	P67	P80	Item	Disc	P50	P67	P80
MT1.2	3	-1.17	-0.93	-0.70	MT2.11	3	-0.04	0.19	0.42
MT1.1	3	-1.14	-0.91	-0.68	MT2.19	5	0.03	0.17	0.31
MT1.13	5	-0.96	-0.82	-0.69	MT2.24	4	0.04	0.21	0.38
MT1.3	5	-0.93	-0.79	-0.65	MT1.25	4	0.07	0.25	0.42
MT1.21	3	-0.92	-0.69	-0.46	MT2.9	3	0.11	0.35	0.57
MT1.8	5	-0.91	-0.77	-0.63	MT2.8	4	0.12	0.30	0.47
MT1.9	3	-0.84	-0.61	-0.38	MT1.26	2	0.14	0.49	0.83
MT1.18	3	-0.77	-0.54	-0.31	MT1.30	3	0.17	0.41	0.63
MT1.24	4	-0.76	-0.58	-0.41	MT2.12	3	0.18	0.42	0.64
MT1.15	2	-0.74	-0.39	-0.05	MT2.10	4	0.25	0.43	0.60
MT1.12	4	-0.74	-0.56	-0.39	MT2.17	4	0.26	0.43	0.60
MT1.10	1	-0.63	0.08	0.76	MT2.3	3	0.26	0.50	0.72
MT1.17	4	-0.60	-0.42	-0.25	MT1.11	2	0.29	0.64	0.98
MT1.14	4	-0.56	-0.38	-0.21	MT2.16	3	0.30	0.53	0.76
MT1.5	4	-0.49	-0.31	-0.14	MT2.23	3	0.31	0.54	0.77
MT2.1	4	-0.45	-0.28	-0.11	MT2.22	4	0.33	0.50	0.67
MT1.16	4	-0.42	-0.24	-0.07	MT2.27	3	0.37	0.61	0.84
MT2.2	4	-0.40	-0.22	-0.05	MT1.20	2	0.39	0.74	1.08
MT1.19	3	-0.33	-0.10	0.13	MT2.29	3	0.39	0.63	0.86
MT2.26	4	-0.32	-0.15	0.02	MT2.6	3	0.43	0.67	0.90
MT1.6	3	-0.29	-0.05	0.17	MT1.27	3	0.45	0.69	0.91
MT2.4	2	-0.25	0.10	0.44	MT2.15	3	0.51	0.74	0.97
MT1.23	4	-0.19	-0.02	0.15	MT2.25	3	0.56	0.80	1.02
MT1.29	4	-0.15	0.02	0.19	MT2.7	2	0.58	0.93	1.27
MT1.7	2	-0.13	0.22	0.56	MT2.14	2	0.87	1.22	1.56
MT2.30	4	-0.13	0.04	0.21	MT2.18	2	0.89	1.25	1.59
MT2.21	4	-0.12	0.06	0.23	MT2.13	2	0.92	1.27	1.61
MT1.4	2	-0.11	0.24	0.58	MT2.28	2	1.29	1.64	1.98
MT1.22	3	-0.06	0.18	0.41	MT2.20	1	2.19	2.89	3.57
MT1.28	3	-0.05	0.19	0.41	MT2.5	1	2.93	3.64	4.32

APPENDIX H

CORRECTED ITEM-TOTAL CORRELATIONS OF 2006 SST MATHEMATICS ITEMS

	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted		Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted		Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
MT1.1	0.434	0.948	MT1.21	0.468	0.948	MT2.11	0.573	0.948
MT1.2	0.375	0.949	MT1.22	0.552	0.948	MT2.12	0.412	0.948
MT1.3	0.534	0.948	MT1.23	0.577	0.948	MT2.13	0.192	0.949
MT1.4	0.361	0.949	MT1.24	0.519	0.948	MT2.14	0.264	0.949
MT1.5	0.576	0.948	MT1.25	0.599	0.948	MT2.15	0.473	0.948
MT1.6	0.486	0.948	MT1.26	0.435	0.948	MT2.16	0.476	0.948
MT1.7	0.473	0.948	MT1.27	0.444	0.948	MT2.17	0.545	0.948
MT1.8	0.537	0.948	MT1.28	0.583	0.948	MT2.18	0.194	0.949
MT1.9	0.409	0.948	MT1.29	0.600	0.947	MT2.19	0.639	0.947
MT1.10	0.311	0.949	MT1.30	0.499	0.948	MT2.20	0.183	0.949
MT1.11	0.485	0.948	MT2.1	0.611	0.947	MT2.21	0.606	0.947
MT1.12	0.557	0.948	MT2.2	0.620	0.947	MT2.22	0.527	0.948
MT1.13	0.502	0.948	MT2.3	0.502	0.948	MT2.23	0.449	0.948
MT1.14	0.575	0.948	MT2.4	0.399	0.949	MT2.24	0.533	0.948
MT1.15	0.402	0.949	MT2.5	-0.064	0.950	MT2.25	0.333	0.949
MT1.16	0.640	0.947	MT2.6	0.449	0.948	MT2.26	0.584	0.948
MT1.17	0.586	0.948	MT2.7	0.376	0.948	MT2.27	0.413	0.948
MT1.18	0.439	0.948	MT2.8	0.529	0.948	MT2.28	0.257	0.949
MT1.19	0.481	0.948	MT2.9	0.541	0.948	MT2.29	0.478	0.948
MT1.20	0.443	0.948	MT2.10	0.563	0.948	MT2.30	0.586	0.948

APPENDIX I

DESCRIPTIVE STATISTICS OF 2006 SST ITEM GROUPS

Groups	N	Min	Max.	Mean	Average p	Std.	Skewness	Kurtosis
Basic Calculations	872 956	-1.00	4.00	2.66	0.67	1.47	-0.86	-0.42
Symbolic Calculations I	872 956	-1.75	7.00	2.60	0.37	2.32	0.21	-1.02
Generalizations	872 956	-0.75	3.00	1.18	0.39	1.06	0.03	-0.87
Word Problems	872 956	-1.50	6.00	2.99	0.50	2.19	-0.08	-1.26
Symbolic Calculations II	872 956	-0.50	2.00	0.78	0.39	0.95	0.24	-1.59
Advanced Calculations I	872 956	-1.75	7.00	1.16	0.17	1.88	1.40	1.33
Advanced Calculations II	872 956	-2.50	10.00	0.95	0.10	2.09	1.77	2.90
Advanced Generalizations	872 956	-0.50	2.00	0.10	0.05	0.48	1.40	2.16
Triangle Calculations	872 956	-1.25	5.00	1.17	0.23	1.38	0.95	0.52
Quadrangle Calculations	872 956	-0.75	3.00	0.69	0.23	1.04	0.86	-0.31
Circle Calculations	872 956	-1.25	5.00	1.20	0.24	1.33	0.84	0.10
Analytic Geometry Calculations	872 956	-0.50	2.00	0.39	0.20	0.77	1.08	-0.07
Geometry Problems	872 956	-1.00	4.00	0.59	0.15	1.20	1.51	1.40
TOTAL	872 956	-13.50	60.00	16.46	0.27	13.90	0.80	-0.13

APPENDIX J

ITEM CLASSIFICATION GUIDELINE

Task: Grouping 2006 SST Math Items according to following procedure

1) Read items without solving them and decide what might be measured with that item in terms of cognitive process and content
2) Each item should match with a group
3) Some groups may not have items
4) Write number of question to a related cell.

	Computation -Knowing	Comprehension-Knowing	Application-Applying	Analysis-Reasoning
Basic				
Advanced				
Geometry				

CONTENT

Basic: Content of high school mathematics in which all students cover (before trigonometry topic)

Advanced: Content of high school mathematics TM and Science students cover (after trigonometry topic)

Geometry: Content of Geometry

COGNITIVE PROCESS (According to Bloom)

Computation: Knowledge of specific facts, Knowledge of terminology, Ability to carry out algorithms

Comprehension: Knowledge of concepts, Knowledge of principles, rules and generalizations, Knowledge of mathematical structure, Ability to transform problem elements from one mode to another, Ability to follow a line of reasoning, Ability to read and interpret a problem.

Application: Ability to solve routine problems, Ability to make comparisons, Ability to analyze data, Ability to recognize patterns isomorphisms, and symmetries

Analysis: Ability to solve nonroutine problems, Ability to discover relationships, Ability to construct proofs, Ability to criticize proofs, Ability to formulate and validate generalizations.

COGNITIVE PROCESS (According to TIMSS)

Knowing: covers the facts, procedures, and concepts students need to know

Applying: focuses on the ability of students to apply knowledge and conceptual understanding to solve problems or answer questions

Reasoning: goes beyond the solution of routine problems to encompass unfamiliar situations, complex contexts, and multi-step problems.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Arıkan, Serkan
Nationality: Turkish (T.C.)
Date and Place of Birth: 3 September 1978, Aydın
Marital Status: Married
Phone: +90 532 327 00 39
email: serkanarikan@gmail.com

EDUCATION

Degree	Institution	Year of Graduation
MA	The Ohio State Univ. Quantitative Research Evaluation and Measurement	2004
BS	Boğaziçi Univ. Teaching Mathematics	2000
High School	Ortaklar Anatolian Teacher High School, Aydın	1995

WORK EXPERIENCE

Year	Place	Enrollment
2008- Present	Cito Turkey	Science Committee Member
2004 - 2008	ÖSYM	Researcher
2000 - 2002	F.M.V. Private Işık High School	Mathematics Teacher
1999 - 2000	Boğaziçi Univ.	Student Assistant

PUBLICATIONS

1. Demirtaşlı, N. Ç., Arıkan, S. (2008). Madde Tepki Kuramı’nın ÖİS uygulamalarındaki yeri. *Paper presented at National Conference of Measurement and Evaluation on Education and Psychology, Ankara, Turkey, 14-16 May 2008.*

2. Arıkan, S., Konak, Ö. A., Yağmur, K. (2008). Bilişsel Gelişimin ÖİS kapsamında ölçülmesi. *Paper presented at National Conference of Measurement and Evaluation on Education and Psychology, Ankara, Turkey, 14-16 May 2008.*

3. Arıkan, S. (2008). Cito Türkiye Öğrenci İzleme Sisteminde (ÖİS) Bilişsel Gelişim Uygulaması. *Cito Eğitim: Kuram ve Uygulama Dergisi*, Tanıtım Sayısı, pp. 28-32.

4. Berberoğlu, G., Arıkan, S., Demirtaşlı, N., Güzel., Ç. İ., Tuncer, Ç. Ö. (2009). İlköğretim 1.-5. Sınıflar Arasındaki Öğretim Programlarının Kapsam ve Öğrenme Çıktıları Açısından Değerlendirilmesi. *Cito Eğitim: Kuram ve Uygulama Dergisi*, 1, pp. 9-48.

5. Demirtaşlı, N., Berberoğlu, G., İşgüzel., Ç. , Özgen, Ç. Konak, Ö. A., Arıkan, S. (2009). Öğrenci Yeterliklerinin Belirlenmesini amaçlayan bir ölçme değerlendirme sistemi örneği: Cito Türkiye Öğrenci İzleme Sistemi. *Paper presented at National Conference on Turkey's Vision 2023 University Entrance System, Ankara, Turkey, 16-17 April 2009.*

6. Berberoğlu, G., Demirtaşlı, N., Güzel, Ç., Arıkan, S. (2009). Affective Characteristics related to student cognitive development in the Turkish Pupil Monitoring System (PMS). *Paper presented at International ENSEC Conference, İzmir, Turkey, 9-12 September 2009.*

7. Güzel., Ç. İ., Berberoğlu, G., Demirtaşlı, N., Arıkan, S., Tuncer, Ç. Ö. (2009).Öğretim Programlarının Öğrenme Çıktıları Açısından Değerlendirilmesi. *Cito Eğitim: Kuram ve Uygulama Dergisi*, 6, pp. 9-30.