T.C.
BAHÇEŞEHİR ÜNİVERSİTESİ

THE ANALYSIS OF SUBTRACTION OF FRACTIONS THE RULE SPACE METHOD AT GRADE 7

Master's Thesis

H.GÜLAY TEZCAN

İSTANBUL, 2008

T.C.
BAHÇEŞEHİR ÜNİVERSİTESİ
INSTITUTE OF SCIENCE
INFORMATION TECHNOLOGIES PROGRAM

THE ANALYSIS OF SUBTRACTION OF FRACTIONS THE RULE SPACE METHOD AT GRADE 7

Master's Thesis

H.GÜLAY TEZCAN

Thesis Supervisor: ASST.PROF. DR. ORHAN GÖKÇÖL

İSTANBUL, 2008

T.C.
BAHÇEŞEHİR ÜNİVERSİTESİ

DEPARTMENT OF COMPUTER ENGINEERING

INFORMATION TECHNOLOGIES PROGRAM

Name of the Thesis: The Analysis of Subtraction of Fractions Using Rule Space Method At Grade 7
Name Last Name of the Student: H. Gülay Tezcan
Date of Thesis Defense: Eylül 2008

The thesis has been approved by the Institute of Department of Computer Engineering Information Technologies Program

Prof. Dr. A. Bülent ÖZGÜLER
Director

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

Signature
Asst. Prof. Doç. Dr. Orhan GÖKÇÖL
Program Coordinator
Signature

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

Examining Committee Members	Signatures
Asst.Prof.Dr. Orhan GÖKÇÖL Thesis Supervisor	------------------
Asst.Prof.Dr. M. Alper TUNGA Member	------------------
Asst.Prof.Dr. Yalçın ÇEKİÇ Member	------------------

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor Asst.Prof. Dr.Orhan Gokcol , for his valuable guidance and advice. He has been very supportive and patient throughout the progress of my thesis.

Also Iwould like to thank to Prof. Dr. Kumi Tatsuoka, who is invited to Bahcesehir College to give a three days seminar to mathematics department for helping me to gain new ideas.
I would heartily like to thank to, my parents , my friends, my students and the teachers who are mathematics department in Bahcesehir College who all have continiuously supported me spiritually during my studies thesis.

May be this thesis will be useful for the other concerners of this topic.

September, 2008 **H. Gulay TEZCAN**

ABSTRACT

THE ANALYSIS OF SUBTRACTION OF FRACTIONS THE RULE SPACE METHOD AT GRADE 7

Tezcan, H.Gulay

Department of Computer Engineering
Information Technologies Program

Supervisor: Asst.Prof. Dr. Orhan Gokcol

September 2008, 88 Pages

This study is one application of The Rule Space Model (RSM) at 7th grade students in Bahcesehir College. The purpose is to measure the student's performances with the attributes of subtraction of fractions.

First of all, the attributes for the subtraction of fractions are listed. Secondly the attributes of the test questions' are coded in with the attributes in Q matrix form. Then response patterns are obtained using Q-matrix. The Rule-Space Model analyzes the score according to test takers' specific skills; based on the answer choice he or she has made.

For the 119 students, a personal attribute mastery report is prepared which shows how he or she has done individually. This individual report breaks down the total score into categories of mastery and where the student needs improvement. Also class reports which help the teacher to focus on the attributes that may be developed.

The results are between 7 and 20 nets and the average of the class is 16, 92 out of 20 questions. This result can be interpreted as the teacher's success in performing the task well.

Keywords: Attribute, Rule Space Model,Q-matrix,

OZET

7. SINIFLARDA UYGULANAN KESIRLERDE CIKARMA ISLEMI TESTININ RULE SPACE MODELI ILE ANALIZI

Tezcan, H.Gulay

Bilgi Teknolojileri Programı

Tez Danısmanı: Yrd. Doc. Dr. Orhan Gokcol

Eylul, 2008, 88 Sayfa

Bu calısma Rule Space Modeli'nin (RSM) Bahcesehir Koleji 7. sınıf ogrencileri uzerinde bir uygulamasıdır. Amac ogrencilerin performanslarını kesirlerde cıkarma isleminin bilgi basarı bilesenleri (BBB) ile olcmektir.

Her seyden once kesirlerde cıkarma islemi icin bilgi basarı bilesenleri listelendi. Sonra test sorularının bilgi basarı bilesenleri Q-matris formunda kodlandı. Donut sonucları Q-matris kullanılarak belirlenmistir. Rule Space Modeli sonucu analiz ederken testi uygulayanların verdikleri yanıtlara gore ozel becerilerini belirler.

119 ogrenci icin kisisel bilgi beceri bilesenlerindeki hakimiyetlerini iceren bir rapor hazırlanmıstır. Bu bireysel rapor, testi uygulayan ogrencilerin toplam puanını hakimiyet alanlarına ve gelistirilebilir alanlarına gore ayırır. Ayrıca sınıf ogretmeninin agırlık vermesi gereken bilgi beceri bilesenlerini gormesi icin sınıf raporu hazırlanmıstır.

Sonuclar 7 ile 20 net arasında ve sınıfın ortalaması 20 soru uzerinden 16,92 nettir. Bu sonuc ogretmenin konuyu iyi anlattıgı seklinde yorumlanabilir.

Anahtar Kelimeler: Q matrisi(Q matrix),Rule Space Model (Tanisal Test Modeli) , Bilgi beceri bileseni(Attribute)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Classical Theory Test	: CCT
Item Response Theory	: IRT
Knowledge State	: KS
Rule Space Model	: RSM
Scholastic Aptitude Test	: SAT

1. INTRODUCTION

Recent developments in cognitive theories have shown that learning is the reorganization and integration of complex tasks. However, learning models considered by educational measurement are primarily linear, and hence measurement models that have been developed support the view of ability levels. The purpose and goal of these models are focused on making inferences about amount of ability or amount of knowledge that an individual possess, which can be located on the continuum.

A new view of achievement that emerges from cognitive and domain studies emphasizes the importance of how knowledge is organized, what processes are used to solve problems , the degree to which certain procedures are automated, and the ability to represent knowledge in a variety of ways. New measurement models should be able to measure such abilities, as well as traditional ability levels.

From the perspective of more traditional approaches, such as classical test theory, an advantage of Item Response Test (IRT) is that it potentially provides information that enables a researcher to improve the reliability of an assessment. This is achieved through the extraction of more sophisticated information regarding psychometric properties of individual assessment items. IRT is sometimes referred to using the word strong as in strong true score theory or modern as in modern mental test theory because IRT is a more recent body of theory and makes more explicit the hypotheses that are implicit within Classical Test Theory (CTT).

In psychometrics, item response theory (IRT) is a body of theory describing the application of mathematical models to data from questionnaires and tests as a basis for measuring abilities, attitudes, or other variables. It is used for statistical analysis and development of assessments, often for high stakes tests such as the Graduate Record Examination. At its most basic level, it is based on the idea that the probability of getting an item correct is a function of a latent ability. For example, a person with higher intelligence would be more likely to correctly respond to a given item on an intelligence test.Items may be questions that have incorrect and correct responses, statements on questionnaires that allow respondents to indicate level of agreement, or patient symptoms scored present/absent.

In education, Psychometricians apply IRT in order to achieve tasks such as developing and refining exams, maintaining banks of items for exams, and equating for the difficulties of successive versions of exams

IRT is often referred to as latent trait theory, strong true score theory, or modern mental test theory and is distinguished from Classical test theory.

IRT models are used as a basis for statistical estimation of parameters that represent the 'locations' of persons and items on a latent continuum or, more correctly, the magnitude of the latent trait attributable to the persons and items. For example, in attainment testing, estimates may be of the magnitude of a person's ability within a specific domain, such as reading comprehension. Once estimates of relevant parameters have been obtained, statistical tests are usually conducted to gauge the extent to which the parameters predict item responses given the model used. Stated somewhat differently, such tests are used to ascertain the degree to which the model and parameter estimates can account for the structure of and statistical patterns within the response data, either as a whole, or by considering specific subsets of the data such as response vectors pertaining to individual items or persons. This approach permits the central hypothesis represented by a particular model to be subjected to empirical testing, as well as providing information about the psychometric properties of a given assessment, and therefore also the quality overestimates.

One of the major contributions of item response theory is the extension of the concept of reliability. Traditionally, reliability refers to the precision of measurement (i.e., the degree to which measurement is free of error). And traditionally, it is measured using a single index defined in various ways, such as the ratio of true and observed score variance. This index is helpful in characterizing a test's average reliability, for example
Classical test theory is a body of related psychometric theory that predicts outcomes of psychological testing such as the difficulty of items or the ability of test-takers. Generally

speaking, the aim of classical test theory is to understand and improve the reliability of psychological tests.

Classical test theory is by far the most influential theory of test scores in the social sciences. In psychometrics, the theory has been superseded by the more sophisticated models in Item Response Theory (IRT). IRT models, however, are catching on very slowly in mainstream research. One of the main problems causing this is the lack of widely available, user-friendly software; also, IRT is not included in standard statistical packages like SPSS. As long as this problem is not solved, classical test theory will probably remain the theory of choice for many researchers.

At the item level, the CCT model is relatively simple. CTT doesn't invoke a complex theoretical model to relate an examinee's ability to success on a particular item.

(http://en.wikipedia.org/wiki/Classical_test_theory
http://en.wikipedia.org/wiki/Item_response_theory)

1.1. A COMPARISON OF CLASSICAL AND ITEM RESPONSE THEORY

Classical Test Theory (CTT) and Item Response Theory (IRT) are largely concerned with the same problems but are different bodies of theory and therefore entail different methods. Although the two paradigms are generally consistent and complementary, there are a number of points of difference:

IRT makes stronger assumptions than CTT and in many cases provides correspondingly stronger findings; primarily, characterizations of error. Of course, these results only hold when the assumptions of the IRT models are actually met.

Although CTT results have allowed important practical results, the model-based nature of IRT affords many advantages over analogous CTT findings. CTT test scoring procedures have the advantage of being simple to compute (and to explain) whereas IRT scoring generally requires relatively complex estimation procedures (note that in the Rasch model the total score for a person is the sufficient statistic of the person parameter).

IRT provides several improvements in scaling items and people. The specifics depend upon the IRT model, but most models scale the difficulty of items and the ability of people on the same metric. Thus the difficulty of an item and the ability of a person can 3 be meaningfully compared.

Another improvement provided by IRT is that the parameters of IRT models are generally not sample-or-test-dependent whereas true-score is defined in CTT in the context of a specific test. Thus IRT provides significantly greater flexibility in situations where different samples or test forms are used. These IRT findings are foundational for computerized adaptive testing.

It is worth also mentioning some specific similarities between CTT and IRT which help to understand the correspondence between concepts.

While the concept of the item response function has been around since before 1950, the pioneering work of IRT as a theory occurred during the 1950s and 1960s. Two of the pioneers were the Educational Testing Service psychometrician Frederic M. Lord and, independently, the Danish mathematician Georg Rasch.

However, while the mathematical groundwork was laid, IRT did not become widely used until the late 1970s and 1980s when the advent of the personal computer provided the computer power necessary to a greater number of researchers

(http://en.wikipedia.org/wiki/Classical_test_theory
http://en.wikipedia.org/wiki/Item_response_theory
http://ltj.sagepub.com/cgi/content/abstract/15/2/119)

12. THE RULE SPACE METHOD

Test item development has been a theoretical in terms of cognitive theory.(Gitomer 1998). It is important to understand the nature of cognitive processing involved in the subject. Gitomer pointed out that students errors are often linked to an inability to conceptualize a problem, to a failure to employ efficient problem-solving heuristic, and to lack of willingness to pursue difficult problems that cannot be solved quickly. The teachers realized that some students have a view of mathematics that it is simply equivalent to the learning algorithms. However, Gitomer (1988) developed a diagnostic test that was to designed to measure knowledge, execution referred to the procedural evaluation of a problem, application involved in recognizing a procedure to execute for a given problem, decomposition processes that require decomposing a problem with multiple sub goals had a strong relationship with mathematics grades.

With school districts placing more importance on test scores, it's important to know what all of these numbers mean. Using statistics, probability and computer coding, Kikumi Tatsuoka has come up with the Rule-Space Model. Tatsuoka, a full-time project director in TC Department of Human Development, Distinguished Research Scientist and Adjunct Professor of Statistics and Education, developed the model while working at ETS. “Her Rule-Space Model (RSM) is seen as one of the most viable alternatives to the traditional unidimensional models of Item Response Theory (IRT)," said James Corter, Chair of Human Development.

13. DEFINITION OF PROBLEM

The rule space is used to classify students into one of two methods for solving subtraction of fractions and diagnose knowledge states in this study. A 20-item test was administered to 119, 7th grade students in Bahcesehir College.

A set of underlying cognitive processing skills and knowledge believed to be involved in solution of the 7th grade mathematics test (subtraction of fractions) was identified. Then the teachers coded the test items in terms of which attributes are required for successful solution of each item, a process that defines the Q- matrix. After preparation of the data set, The Rule Space analysis was performed using special purpose software developed for this purpose. Results of the RSM include diagnosis of each student in terms of a vector of attribute mastery probabilities, as well as classification of each student into a closest knowledge state. And then the diagnostic scoring reports are given to the students and for the each class teacher, class reports are prepared.

14. THE ROADMAP FOR THE THESIS

Thesis is divided into five chapters. In chapter 1, the important aspects of the measurement techniques and item response theory – classical test theories reviewed. In the second part, the description of the Rule Space Model is given. In section three, the subtraction fraction test prepared and results are discussed. The RSM analyses were run separately for each of the classes in part four. A report for the class teacher and for each class diagnostic scoring reports are prepared. The results are discussed and in section five.

2. THE RULE SPACE METHOD

2.1. WHAT ARE THE GOALS OF THE RULE SPACE METHOD?

The value of diagnostic profile that enumerates strengths and weakness in individual performance has been recognized by several studies. (Tatsuoka & Tatsuoka 1998; Tatsuoka 1995; Van Lehn1981) . However, valuable diagnostic profiles must include the information about how well test takers performed on the underlying knowledge and cognitive processing skills required in answering problems. In the rule space approach, knowledge and cognitive processing skills will be called attributes and binary attribute patterns that express mastery / non-mastery of attributes define knowledge states or latent knowledge states. Attributes are latent variables that are impossible to observe directly. Moreover, the domains of our interest usually involve several to hundreds of attributes.It is a almost impossible to make an inferences on such a large number of unobservable attributes.

In late 1970s and early 1980s, Brown (1979) ,Van Lehn (1981) and other researchers in computer science have programmed intelligent tutoring systems and error diagnostic systems. Especially, Brown and Burton's " Buggy System generated a sensation to the educational community by showing that a computer can diagnose student's procedural errors called "bugs " whole number operations. They have programmed hundreds of bugs discovered by expert teachers in their computer systems, and used this information for building their system for diagnostic bugs. In a similar manner to Buggy system, Tatsuoka & Bailie (1984) have developed a computer program (FBUG System) for diagnosing erroneous rules of operations in fraction addition and subtraction problems. This program required predetermined response patterns obtained from various rules of operations, which have been found previously by teachers or by doing an error analysis on student's responses. However, these approaches were deterministic and impossible to take errors of measurements in considerations. This Rule Space has been developedfor solving this problem. In other words the rule space method is a probabilistic approach specifically developed to deal with the variability of students' responses. Since the rule space is based on a probabilistic approach utilizing the multivariate decision theory, one can compute a membership probability as well attribute mastery probabilities for each student. The methodology is an individualized, statistical pattern classification technique.(Duda & Hart 1973 ; Fukunaga 1990)

2.2 HOW TO MAKE GOOD USE OF THE RULE SPACE PROGRAM?

There are several ways to use the rule space program, for example as a statistical tool for hypothesis testing, and / or finding solution strategies for solving problems, and / or developing cognitive models. There may be creative applications but we try to discuss them below:

Hypothesis Testing:

One of the most important features is that it can be used as a statistical tool for hypothesis testing. The hypotheses to be tested here are experts' opinions and views' on what cognitive skills and knowledge would be used for answering the problems. For example, researchers would like to test their hypothesis that the attribute, "quantities' reading " in the problems such as "at least" ,"two in row or necessary and sufficient conditions "causes frequent errors. They believe this attribute affects the performance of students with the below average more than those in the above average. Then, the researchers can create a response pattern in at test such as SAT Mathematics by coding items involving "quantitative reading "to 1, and others 0 throughout 60 test items. The resulting 60 – item pattern has 1s for the items involving the attribute, and 0s for the items not involving the attribute. Then the compliment of the original pattern implies that the students who don't have this skill are highly likely to get the scores of 0 for the items involving this attribute. Thus, the compliments pattern can be considered as a knowledge state characterized by"can't do only quantitative reading ".

Similarly one can create several response patterns resulting from the lack of some other skills, which the researchers would like to test. Then the rule space program III will provide them with a list of students whose response patterns are very close to the input information with their total scores, and their estimated IRT θ values. The program also gives the membership probability for this knowledge state, along with the mastery probability of the skills the researchers want to test. Therefore, the researchers can use this information for creating their models or cognitive models. In psychometrics and educational measurement, one of the best known examples of a statistical modeling approach is item response theory (IRT). Many models of test performance assume some kind of algebraic relationship on a talent variable (or a few latent variables) to explain observed responses. In IRT models, logistic functions are used on the latent variable θ (ability) to explain students' item responses. The latent variable θ is viewed as ability, or a trait to perform well on test items. In a statistical modeling approach, it is crucial to test how well the model fits the observed responses. Various fit statistics have been developed for this purpose (Glas & Meijer 2003).

Finding students' underlying strategies for solving problems and constructing hierarchical relationships among skills involved in a domain:

Researchers create a flow chart of describing strategies for solving fraction addition and subtraction problems. And there are two methods to solve these questions.

Method A converts a mixed or whole number to an improper fraction and adding or subtracting two fractions. This method is especially noticed in subtraction problems where borrowing is needed. However, since the Method A involves with manipulation with a larger numbers it is more subjected to arithmetic errors.

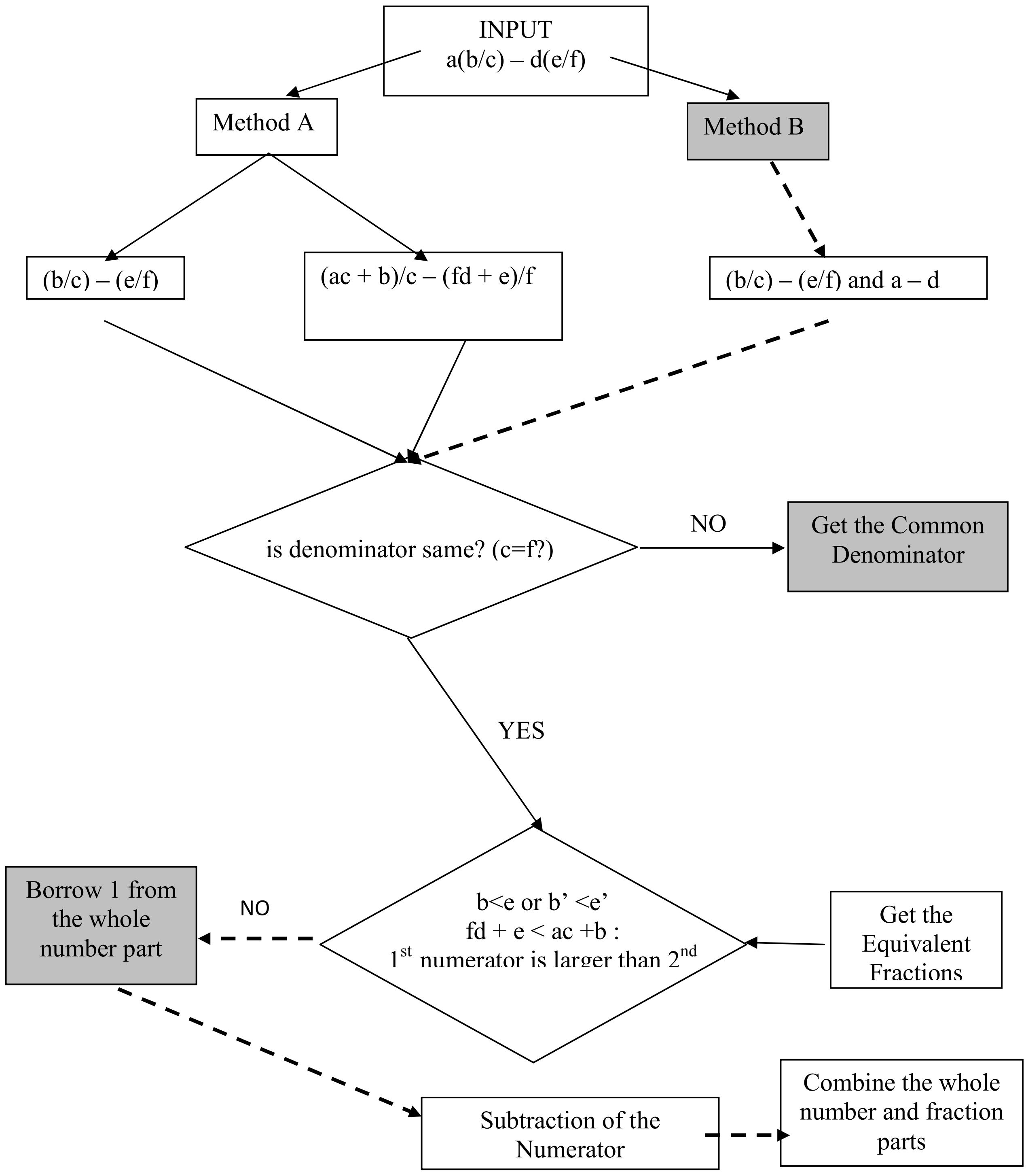


Figure 2.1: A Flow Chart of the Fraction Subtraction Operation for (a (b/c) – d (e/f))

Source: GUIDE TO USING THE RULE SPACE PROGRAM

Method B, separates the fraction part and subtraction part. The advantage of this method is that once the whole number is separated, the student can manipulate smaller numbers.

However, when using this method the student has to remember to repeat the operation twice: once with the whole number part and then with the fraction part.

When the students use Method A, then they don't have to know the borrowing operation. So, they can subtract the numerators without borrowing 1 from the whole part. However, if students use Method B and have not mastered the borrowing operation, then they make errors, with a very high probability, for the items requiring borrowing. Therefore, the items requiring borrowing can be used to find which method students are likely using. As for input information to the program, Rule Space III, one can create an item pattern consisting of 1 non- borrowing item and 0 for borrowing items. Treat this item pattern as the knowledge state, "lack of the borrowing the skill but all other skills can do " for Method B users. Similarly one can create an item pattern of " lack of the column borrowing skill for Method A users. Then the researchers can create item patterns corresponding to the knowledge state, " Lack of skills for getting the common denominators but can do all other skills", or " lack of skills in converting a whole number to a fraction " so and so forth. By creating item patterns systematically by assuming the lack and mastery of various combinations of skills from the flowchart, the researchers can come up with many knowledge states that can be tested with the Rule Space III program. Rule Space analysis results will give the mastery probability of each skill as the membership probability of each individual and each knowledge state from which the individual comes with high probability values. By analyzing the probability values, one can construct a hierarchical relationship among the skills that the researchers tested, whichever / Method A or B users (Kim and Tatsuoka 1995)

<u>Spotting aberrant response patterns resulting from " bugs " or " not appropriate performance on a test :</u>

Rule space technique applicable to a general domain for spotting aberrant response patterns that are derived from erroneous rules of operation, or / and from the lack of a skill or several combinations of skills (Tatsuoka & Linn 1983).Of course the Buggy Systems or expert systems can do the same thing to spot such response patterns, but they require a laborious task for programming. The rule space III program can diagnose such response patterns deviated by few slips from the exact patterns associated with various erroneous rules of operations. The 69 erroneous rules of operations in fraction addition and 54 rules in subtraction problems were discovered by extensive error analyses and interview methods, and these rules were expressed by unique item response patterns of 39-and40 item tests, respectively. Then the rule space was used to diagnose these response patterns, and the studies confirmed that the method worked perfectly well. A few thousands of students response patterns were diagnosed by the method, and the diagnosed results were confirmed by error analyses and interviews (Show, Stanford, Klein & Tatsuoka 1982;Tatsuoka 1984b). The rule space method can be used to diagnose students' errors that were known prior to analysis.

<u>Use the program as a tool for searching unknown or /and unware skills and knowledge</u>

The program can be used to explore unknown sources of errors by retrieving the response patterns clustering to a swarm in the rule space model. Buck and Tatsuoka (1998) found such a swarm at the lower part of low ability section of the rule space in an English listening test, and yet could not be identified by the existing knowledge states. By retrieving response patterns and referring to the test items, they have discovered a new knowledge state characterized by a new attribute (unfamiliar names) in the test. Similarly Tatsuoka (1995) discovered a few knowledge states committed by high ability students in SAT Mathematics. These knowledge states were characterized by the interaction of two mathematical thinking skills, "deductive thinking skills" and "translation of word problems in to equations after identifying variables . By repeating the cycle of testing hypothesis and recovering a swarm that has not been diagnosed, one can easily diagnose all test takers, and develop a cognitive model or instructional model useful in education.

Automatic generation of knowledge states based on experts' hypoteses

However, an error analysis and task analysis is tremendously laborious and time consuming. Tatsuoka (1991) has developed a new technique utilizing Boolean algebra to generate a list of item response patterns corresponding to psychological meaningful knowledge states. Suppose we have identified 17 attributes for a domain of our interest. Then the profiles of mastery / non mastery on the 17attributes have the combination of

$$2^{17} = 131.072 \quad (2.1)$$

patterns which a very large number. Firstly we have to have a clever idea to manage a computational problem for dealing with enormously large numbers of knowledge states in classification because the traditional statistical decision theory has been handling the cases of only a smaller number of classification groups. Secondly, diagnostic reports have to be comprehended easily by test users. They must be simple and clear. So, we have to select the most valid and useful information from enormous volume of resources.

Automated system for classification of individuals to their true latent groups

An intelligent error diagnostic system of (BuggySystem) of whole number subtraction problems developed by Brown and Burton (1979) required predetermined "bugs , identified through thousands of hours of work by trained teachers. (Van Lehn 1981). Once an automated system is build, then error diagnosis can be done automatically for a stream of new examples. There are many such examples in real life such as "classifying an X-ray image of a tumor cancerous or benign , "have recognize hand written often performed by human experts, but it is increasingly becoming feasible to design automated systems to replace the human experts and perform either better or as well as experts do (Riply 996). However, any automated systems and statistical methods that have been currently availed require a train the coefficients in the systems.

Generating a training set

More recently, Neural Networks have risen from analogies with models of the way that the student might approach to acquire new knowledge and skills. This approach has had a great impact on the practice of pattern recognition in engineering and science. The trained neural network can perform either better or equally well to human experts. Hayashi and Tatsuoka (2000) trained a neural network using the rule space results in a fraction subtraction domain.(Klein, Birenbaum, Standiford & Tatsuoka 1982) and confirmed that the trained neural network worked very well for diagnosing fraction errors committed by individual test takers. Mislevy (1995) applied the rule space results from the same fraction test and successfully developed a Bayesian Influence Network. In these studies, the rule space analyses are avaible, and then we can use either neural network or Bayesian Influence Network that are already available in statistics to replace human decisions making processes.

They have a question " if a training set is so valuable, then is it possible to develop it without going through laborious and painstaking task analysis ? The rule space method developed by Tatsuoka and her associates (1989,1990 1991,1993,1995, and 1998, 2000, in press) is designed to generate a training set without relying on human's judgment. The methodology adapts one of the statistical approaches used in the statistical pattern recognition and classification problems. (Fukunaga 1990 ; Duda, R.D.& Hart,P.E. 1973; Ripley 1996).Statistical pattern recognition and classification theory is known as a methodology for assignment of a physical objects or event to one several predetermined categories. Having a computer recognize handwritten letters, or describe the shape of a figure has been one of the most popular applications of this theory.

The rule space method is an extension of the pattern recognition and classification approach developed for diagnosing cognitive errors. While physical objects or events in science applications are observable, attributes and knowledge states are not observable. Measurement of unobservable latent variables can be assessed only indirectly from observable item scores by making inferences about what misconceptions, leading to what incorrect responses, did this subject most likely have. Given a response pattern, we want to determine a knowledge state what is the probability that a test taker's observed responses have been drawn from.

Attributes derived by domain experts are always multidimensional. Indeed, the previous studies of several large-scale assessments such as SAT, TOEFL, TOEIC, and GRE. Quantitative tests have indicated that their attributes are independent by factor analysis. Therefore, a space spanned by attributes or, equally to say, a set of knowledge states is multidimensional. Regarding making inferences about unobservable variables from item scores, Item Response Theory is one of the simplest cases to make an inference about "latent IRT ability measure θ from item responses, because IRT assumes underlying cognitive attributes are undimensional θ. It is realistic that a model or statistical method must be able to handle multi-dimensional latent abilities characterized by many unobservable attributes.

2.3THE DESCRIPTION OF THE METHODOLOGY

The methodology consists of two stages:

1) The phase of extracting cognitive features of a domain and then determining cognitive features of a domain and then determining classification groups.

2) To classify student's response patterns into one of the predetermined knowledge states.

The classified group is, with a high probability value, considered as the individual's latent knowledge state. Also attribute mastery probabilities test takers are computed.

Stage 1

<u>1) Phase of extracting features and making an item- attribute matrix Q</u>

Identification of a set of effective attributes requires statistical hypotheses testing. Firstly, the methodology assumes domain experts to generate their hypotheses on attributes and secondly, examines whether or not their hypotheses are statistically sound.

Having specified attributes, a $k \times k$ incidence matrix Q , in which the rows represent attributes(i.e., tasks, subtasks, cognitive processes and skill, etc.) and columns represent items. The entries in each column indicate which attributes are involved in the solution of each item. (K. K. Tatsuoka 1990)

Example 1. Let us consider the 3×5 incidence matrix, where I_1, I_2 ,I_3 ,I_4 ,I_5, are items A_1, A_2, A_3 are attributes :

$$Q = \begin{array}{c|ccccc} & I_1 & I_2 & I_3 & I_4 & I_5 \\ A_1 & 1 & 0 & 1 & 0 & 1 \\ A_2 & 0 & 1 & 0 & 0 & 1 \\ A_3 & 1 & 1 & 1 & 1 & 1 \end{array} \qquad (2.1)$$

There are three row vectors,

$$A_1 = (\,1\ \ 0\ \ 1\ \ 0\ \ 1\,) \qquad (2.2)$$

$$A_2 = (\,0\ \ 1\ \ 0\ \ 0\ \ 1\,)$$

$$A_3 = (\,1\ \ 1\ \ 1\ \ 1\ \ 1\,)$$

We can say,

Attribute A_1 is involved in Items 1, 3 and 5

Attribute A_2 is involved in Items 2 and 5

Attribute A_3 is in Items 2, 3, 4 and 5

Let us denote these involvement relationships in set- theoretic notation as:

$$A_1 = \{1, 3, 5\} \qquad (2.3)$$

$$A_2 = \{2, 5\}$$

$$A_3 = \{2, 3, 4, 5\}$$

Then, the union set $A_1 \cup A_2$, of A_1 and A_2, is the set {1, 2, 3, 5} of items involving A_1 and A_2. (Tatsuoka and Tatsuoka 1987,1989).

2) Determinationof all possible knowledge stages

Boolen Algebra L_A and L_1 Generated From attributes Vectors And Item Vectors in a Q- Matrix

The theory of Q-matrices conceptually includes both the lattice spaces of attributes and knowledge states, because the set of sets of row vectors formulate a Boolean algebra. Moreover, the two lattice spaces are mathematically connected through the Q-matrix. Each knowledge state represents which attributes an individual has mastered, and testing for mastery or no mastery of an attribute can be done by examining a question or an equivalent class of questions that involves this attribute. (Falmagne and Doignon, 1988) Suppose L_A and L_1 are lattice spaces generated from K attribute vectors and n item vectors, respectively. Because the following discussions about L_A hold equivalently for L_1, only the attribute lattice space is considered. K. K. Tatsuoka (1991) showed that the lattice space L_A satisfies not only the conditions to be lattice but also satisfies the condition to be a Boolean Algebra.

Example 2 : Addition and multiplication operations in L_A are defined by element wise Boolean addition and multiplication of ϕ and I , which are the elements of attribute vectors, A_k and A_n: $A + B = A \cup B$ and $A \times B = A \cap B$. (2.4)

So we remember the 3×5 incidence matrix, where I_1, I_2, I_3, I_4, I_5, are items A_1, A_2, A_3 are attributes, and obtain Q-matrix

$$Q = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} \text{ so } \phi = \sum_{k=1}^{K} A_k = (0\ 0\ 0\ 0\ 1) \text{ and } I = \sum_{k=1}^{K} A_k = (1\ 1\ 1\ 1\ 1) \qquad (2.5)$$

So $\quad A'_1 = (0\ \ 1\ \ 0\ \ 1\ \ 1\ )$

$A'_2 = (1\ \ 0\ \ 1\ \ 1\ \ 1\ )$

$A'_3 = (1\ \ 0\ \ 0\ \ 0\ \ 1\ )$

Boolean Algebra has an advantage over lattices, because the former is more convenient computationally. The framework of Q-matrix enables the use of a mapping function. There are interesting relationships among the attribute vectors that are totally ordered.

If attribute vector A_k is larger than or equal to A_l, $A_k \geq A_l$, and

A_l is larger than A_m, $A_l \geq A_m$,

that is, $A_k \geq A_l \geq A_m$, then their sum equals the largest attribute, and their product equals the smallest attribute:

$$A_k + A_l + A_m = A_k \quad \text{and} \quad A_k \times A_l \times A_m = A_m$$

These properties are called degenerative relations (Birkoff 1970) and the have an important role for reducing the number of knowledge states.(Varadi & K.K. Tatsuoka 1989)

Stage 2

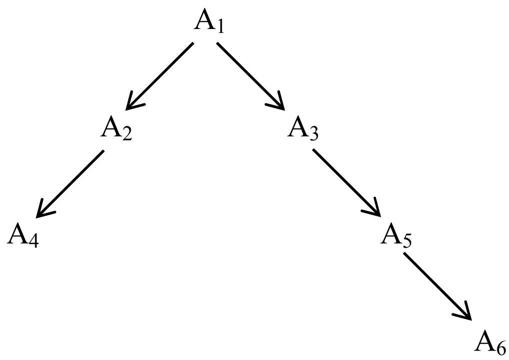
1Prerequisite Relationship Among Attributes: Adjacency Matrix

Boolean algebra is a lattice with the distributive law and partial order of two vectors by an inclusion relationship. On the other hand, the prerequisite relationship, A_k is an immediate prerequisite of A_k (A_k is required for mastering A_l) , originates from the cognitive demands unique to a content domain.(Sato,1990)

From A_k to A_l .The direct relations among the attributes can be represented by a matrix called the adjaceny matrix with elements

$$Q_{jk} = \begin{cases} 1 \text{ if a direct relation exists } A_k \text{ to } A_1 \\ 0 \text{ otherwise} \end{cases}$$

Example: Suppose five attributes A_1, A_2, A_3, A_4 and A_5 are related as follows:



Then, the adjaceny matrix is given by:

$$Q = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \end{array} \begin{array}{c} \begin{array}{cccccc} A_1 & A_2 & A_3 & A_4 & A_5 & A_6 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

2Reachability Matrix

The reachable matrix R is given by computing the powers of (Q + I) with respect to Boolean operations until the result becomes invariant.

It means $(Q + I)^2 = (Q + I)^3 = (Q + I)^4$ …… and R is given by $(Q + I)^2$

Example :

$$R = \left[\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}\right]^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}^2$$

3Definition of Sufficient Item Pool

The item pool associated with a sufficient Q-matrix is called a sufficient item pool. In other words, this condition will navigate processes of item construction and test design toward achieving the goal of measuring objectives. (K. K. Tatsuoka 1993).

As a result the construct validity of a test will improve. It is important to note that a sufficient Q-matrix is the core of a knowledge structure, and the partial order induced by the inclusion relation makes the Q-matrix theory capable of representing the prerequisite relationships among the attributes.

Example: A sample of the Q-matrix of fraction addition problems (K.K. Tatsuoka & M.M. Tatsuoka 1982) is expressed by the influence diagram. The pair wise comparison of the row vectors in a Q-matrix leads to a reachability matrix if the Q-matrix is sufficient. Similarly, the pair wise comparison of the column vectors leads to an item tree. (K. K. Tatsuoka 1990)

Subject: The addition of fractions

i) Method A:

a(b/c) + d(e/f) = (ac + b)/ c + (df +e) / f,

The common denominator l = mc, l = nf ; m and n are integers

= [m (ac+b)]/l + [n(df + e)/ l]

= [(mac +mb) (ndf + ne)] / l

ii) Attributes:

A1	Convert the first mixed number to a simple fraction
A2	Convert the second mixed number to a simple fraction
A3	Take the common denominator and make equivalent fractions
A4	Add the two numerators
A5	Answer the to be simplified to the simplest term

iii) Items:

1) $2\frac{8}{6}+3\frac{10}{6}$ 2) $\frac{2}{5}+\frac{12}{8}$ 3) $\frac{3}{9}+\frac{6}{9}$

4) $2\frac{1}{2}+4\frac{2}{4}$ 5) $\frac{1}{2}+\frac{1}{7}$ 6) $\frac{1}{3}+\frac{1}{2}$

7) $\frac{2}{5}+\frac{1}{5}$ 8) $3\frac{1}{6}+2\frac{3}{4}$ 9) $1\frac{2}{5}+\frac{3}{5}$

iv) The incidence matrix Q:

Table 2.1: The incidence of matrix Q

	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9
A1	1	0	0	1	0	0	0	1	0
A2	1	0	0	1	1	0	0	1	1
A3	0	1	0	1	1	1	0	1	0
A4	1	1	1	1	1	0	0	0	0
A5	1	1	1	1	1	1	1	1	1

v) The influence diagram for nine fraction addition problems

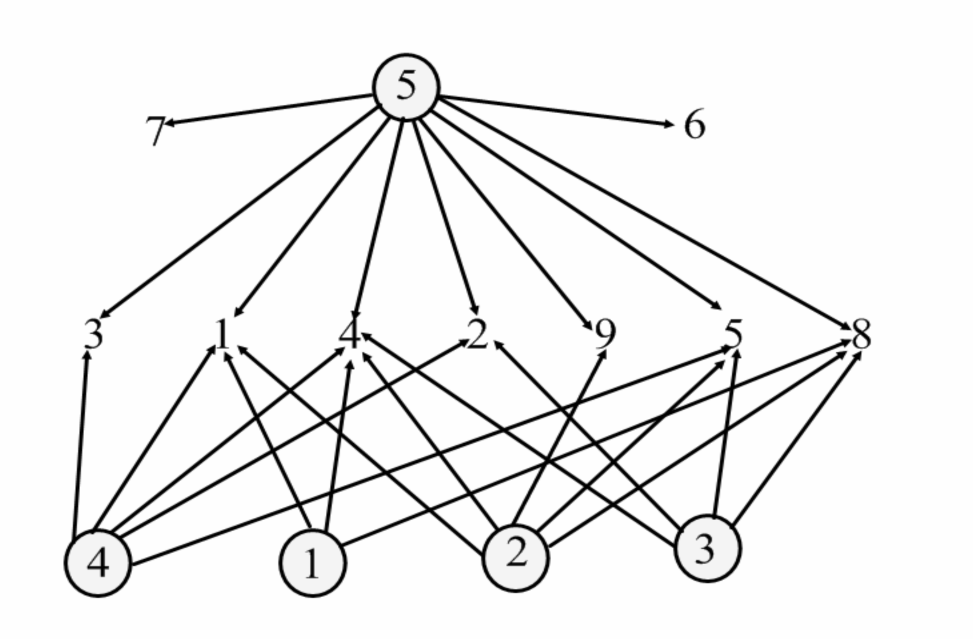


Figure 2.1: The influence diagram for nine fraction addition problems

3. USE OF RULE SPACE MODELLING TO 7TH GRADE STUDENTS

3.1. A GENERAL INFORMATION ABOUT THE TEST AND THE GROUP

The present analyses use the rule space methodology (K.Tatsuoka 1983, 1985,1990, 1995, 1997, in press; K. Tatsuoka & M.Tatsuoka 1987; M.Tatsuoka & K. Tatsuoka 1989) to diagnose each student in terms of mastery of specific "attributes (knowledge and sub skill components) assumed to underlie test performance. This work followed the general outline of any rule space analysis, as follows:

First, in order to identify the specific knowledge and sub skills attributes required for seventh grade subtraction of fractions test items.

Then, with the teachers coded the test items in terms of which attributes are required for successful solution of each item, a process that defines the Q-matrix.

After preparation of data set this is 119 Bahcesehir 7th grade students from 5 classes, the rule space analysis was performed using purpose software developed for this purpose. Results of the RSM include diagnosis of each student in terms of a vector of attribute mastery probabilities, as well as classification each student into a closest knowledge state. Now we can see the details as follow.

3.2.ATTRIBUTE S, BASED ITEMS AND Q– MATRIX

Attributes are defined to be skills at the level representing the scores of misconceptions such as "getting the common denominator ", or "simplify a mixed number before subtraction .

Examples of coding:

Let us start from the first attribute

A1 (Convert a whole number to a fraction or a mixed number).

Suppose an item $3-2\frac{1}{3}$.

A whole number 3 can be rewritten by, $\frac{9}{3}$, $2\frac{3}{3}$ or $1\frac{6}{3}$ .This is A1.

In order to subtract the second fraction, $2\frac{3}{3}$ is the right choice. Then we have to subtract whole number part and fraction part separately, $(2-2)+\left(\frac{3}{3}-\frac{1}{3}\right)$.

$$3-2\frac{1}{3}=2\frac{3}{3}-2\frac{1}{3}$$

A1 (Convert a whole number to a fraction or a mixed number).

$$2\frac{3}{3}-2\frac{1}{3}=\left(2+\frac{3}{3}\right)-\left(2+\frac{1}{3}\right)=(2-2)+\left(\frac{3}{3}-\frac{1}{3}\right)$$

A2 (Separate a whole number from a fraction)

$$(2-2)+\left(\frac{3}{3}-\frac{1}{3}\right)=0+\frac{2}{3}=\frac{2}{3}$$

A7 (Subtracting the second numerator from the first)

So, the item $3-2\frac{1}{3}$ requires attributes A1, A2 and A7.

Let us consider an item $3\frac{3}{8}-2\frac{5}{6}$.

By taking the common denominator and make equivalent fractions,

we get $3\frac{9}{24}-2\frac{20}{24}$,

$$3\frac{3}{8}-2\frac{5}{6} = 3\frac{9}{24}-2\frac{20}{24}$$

A4 (Find a common denominator and make equivalent fractions)

$$3\frac{9}{24}-2\frac{20}{24}=2\frac{33}{24}-2\frac{20}{24}$$

A5 (Borrow 1 from a whole number part and add the denominator to the numerator)

$$2\frac{33}{24}-2\frac{20}{24}=(2-2)+\left(\frac{33}{24}-\frac{20}{24}\right)=\frac{13}{24}$$

A2(Separate a whole number from a fraction) and

A7(Subtracting the second numerator from the first)

Table 3.1 lists the items we used in this study and Table 2.2 and 2.3 summarize attributes and their involvement in the items.

Table 3.1: The Items in the fraction subtraction test

1) $\frac{5}{3}-\frac{3}{4}=?$	11) $4\frac{1}{3}-2\frac{4}{3}=?$
2) $\frac{3}{4}-\frac{3}{8}=?$	12) $\frac{11}{8}-\frac{1}{8}=?$
3) $\frac{5}{6}-\frac{1}{9}=?$	13) $3\frac{3}{8}-2\frac{5}{6}=?$
4) $4\frac{3}{5}-3\frac{4}{10}=?$	14) $3\frac{4}{5}-3\frac{2}{5}=?$
5) $3\frac{1}{2}-2\frac{3}{2}=?$	15) $2-\frac{1}{3}=?$
6) $\frac{6}{7}-\frac{4}{7}=?$	16) $4\frac{5}{7}-1\frac{4}{7}=?$
7) $3-2\frac{1}{5}=?$	17) $7\frac{3}{5}-\frac{4}{5}=?$
8) $\frac{3}{2}-\frac{3}{2}=?$	18) $4\frac{1}{10}-2\frac{8}{10}=?$
9) $3\frac{7}{8}-2=?$	19) $4-1\frac{4}{3}=?$
10) $4\frac{4}{12}-2\frac{7}{12}=?$	20) $4\frac{1}{3}-1\frac{5}{3}=?$

Table 3.2 is a list of attribute involvement in 20 items. We express this relationship in a matrix form and call it a Q-matrix in two ways.

Table 3.2: A Q-matrix of the fraction subtraction items

Item	Associated Attributes
1	4 , 6 , 7
2	4 , 7
3	4 , 7
4	2 ,3 , 5 , 7
5	2 , 4 , 7 OR 2 , 3 , 7
6	7
7	1 , 2 , 7
8	7
9	2 , 7
10	2 , 5 , 7
11	2 , 5 , 7
12	7
13	2 , 4 , 5 , 7
14	2 , 7
15	1 , 5
16	2 , 7
17	2 , 5 , 7
18	2 , 5 , 6 ,7 OR 1 , 2 , 6 , 7
19	1 , 2 , 3 , 5, 7
20	2 , 3 , 5 , 7

Note: Items 5 and 18 involve two strategies.

Table 3.2: A Q-matrix of the fraction subtraction items

Item	Associated Attributes					
	A1	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**

Item	A1	A2	A3	A4	A5	A6	A7
1	0	0	0	1	0	1	1
2	0	0	0	1	0	0	1
3	0	0	0	1	0	0	1
4	0	1	1	0	1	0	1
5	0	1	1	0	0	0	1
6	0	0	0	0	0	0	1
7	1	1	0	0	0	0	1
8	0	0	0	0	0	0	1
9	0	1	0	0	0	0	1
10	0	1	0	0	1	0	1
11	0	1	0	0	1	0	1
12	0	0	0	0	0	0	1
13	0	1	0	1	1	0	1
14	0	1	0	0	0	0	1
15	1	0	0	0	0	0	1
16	0	1	0	0	0	0	1
17	0	1	0	0	1	0	1
18	1	0	0	0	1	1	1
19	1	1	1	0	1	0	1
20	0	1	1	0	1	0	1

Table 3.3 is the summary of attributes we used in this study.

Table 3.3: List of Attributes in Fraction Subtraction Test

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

3.3.ANALYSIS OF KNOWLEDGE STATES

The total of 119 Bahcesehir Koleji students from 5 classes participated in this study. Firstly the Q-matrix was used for generating all possible knowledge states. Knowledge states are associated with a profile of mastered or non-mastered attributes. Profiles are given below in Table 4 . For instance , for students in State1 (denote as KS1),all the attributes are mastered (1,1,1,1,1,1,1). KS15 has a vector of(1,1,1,0,1,1,1) in Table 2 below, and this vector means that a student in that state cannot do A4 but can do all the other attributes. State 11 (KS11) means one can do A7, A2, A1 and A4 but cannot do A5, A3, or A6.

Table 3.4: A list of possible knowledge states generated from Q-matrix

States	A7	A2	A5	A1	A3	A6	A4
1	1	1	1	1	1	1	1
2	1	1	1	0	1	1	1
3	1	1	1	1	0	1	1
8	1	1	1	1	1	0	1
15	1	1	1	1	1	1	0
4	1	1	0	1	0	1	1
5	1	1	1	0	0	1	1
9	1	1	1	0	1	0	1
10	1	1	1	1	0	0	1
16	1	1	1	1	1	0	0
17	1	1	1	0	1	1	0
19	1	1	0	1	1	1	0
22	1	1	1	1	0	1	0
6	1	1	0	0	0	1	1
11	1	1	0	1	0	0	1
12	1	1	1	0	0	0	1
18	1	1	1	0	1	0	0
20	1	1	0	1	1	0	0
23	1	1	1	1	0	0	0
24	1	1	0	1	0	1	0
26	1	1	1	0	0	1	0
7	1	0	0	0	0	1	1
13	1	1	0	0	0	0	1
21	1	1	0	0	1	0	0
25	1	1	0	1	0	0	0
27	1	1	1	0	0	0	0
14	1	0	0	0	0	0	1
28	1	1	0	0	0	0	0
29	1	0	0	0	0	0	0
30	0	0	0	0	0	0	0

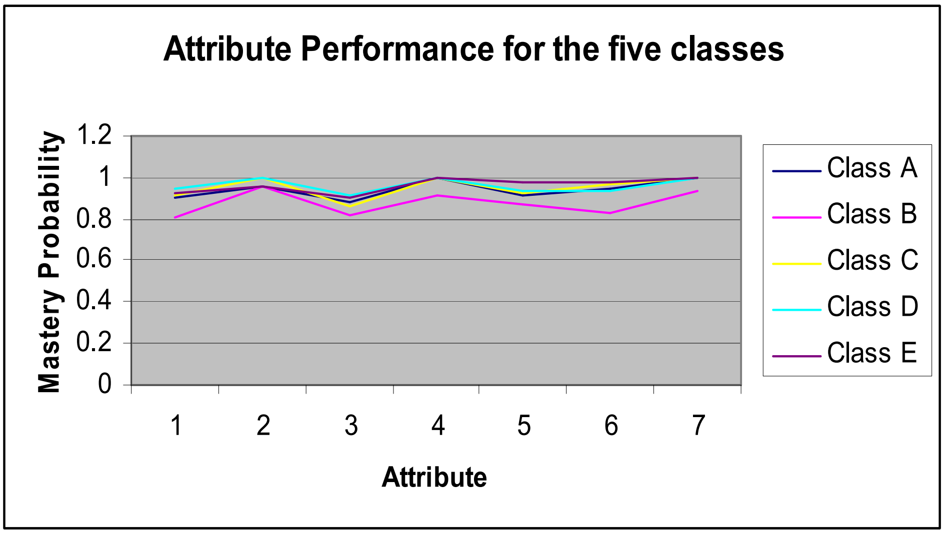


Figure 3.1: Attribute performance for the five classes

Table 3.5: Means of seven attributes across five classes based on linear testing

	A1	A2	A3	A4	A5	A6	A7
Class A	0.90	0.96	0.88	1.00	0.91	0.94	1.00
Class B	0.81	0.96	0.82	0.91	0.87	0.83	0.93
Class C	0.91	1.00	0.86	1.00	0.92	0.97	1.00
Class D	0.94	1.00	0.91	1.00	0.93	0.93	1.00
Class E	0.92	0.96	0.90	1.00	0.93	0.98	1.00

Table 3.5 shows seven attribute means over five classes and their plotting is given in figure 2. Class B performed significantly lower than ($p < .05$) Class C and Class E on attribute A6. There were no other statistically different results among the five classes on any of the other attributes. Student performance on all attributes was generally as more than 75 percent of students are classified in the complete mastery in the complete mastery state of KS1.

Table 3.6: The possible knowledge states of 7th grade students out **of 119**

Number of KS	Number of students	(%)	Mastered Attributes
1	91	76.5	A1, A2, A3, A4, A5, A6, A7
2	2	1.7	A2, A3, A4, A5, A6, A7 (not A1)
3	8	6.7	A1, A2, A4, A5, A6 , A7 (not A3)
4	3	2.5	A1, A2, A4, A6, A7 (not A3 and A5)
5	1	0.8	A2, A4, A5, A6, A7 (not A1 and A3)
6	5	4.2	A2, A3, A4, A6, A7 (not A1 and A5)
7	2	1.7	A4, A6, A7
8	3	2.5	A1, A2, A3, A4, A5, A7 (not A6)
13	1	0.8	A2, A4, A7
17	1	0.8	A2, A3, A5, A6, A7 (notA1 and A4)
30	2	1.7	------

3.4 PERFORMANCE OF ADAPTIVE TESTING AND THEIR COMPARISON TO LINEAR TESTING COMPARISON OF THE MEANS

Table 3.7 lists seven attribute means and standard deviations that were obtained by linear testing and adaptive testing.

Table 3.7: Attribute means of linear and adaptive tests (N = 119)

Linear Testing

	Mean	Standard Deviation
A1	0.89	0.21
A2	0.98	0.13
A3	0.87	0.25
A4	0.98	0.11
A5	0.91	0.25
A6	0.93	0.18
A7	0.99	0.12

Adaptive Testing

	Mean	Standard Deviation
A1_a	0.86	0.21
A2_a	0.97	0.15
A3_a	0.87	0.25
A4_a	0.97	0.14
A5_a	0.90	0.27
A6_a	0.89	0.14
A7_a	0.98	0.12

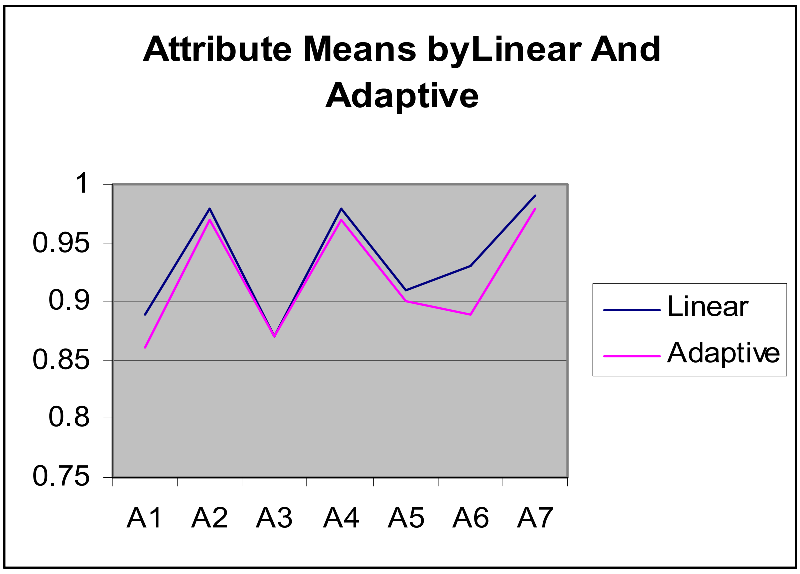


Figure 3.2: Comparison of seven attribute means in linear and adaptive testing

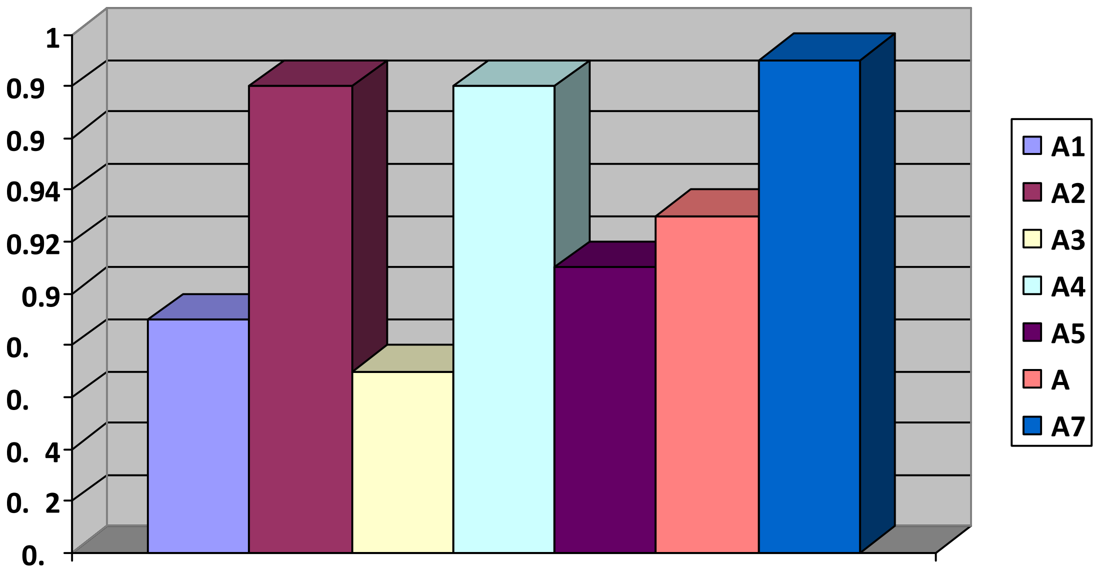


Figure 3.3: Comparison of seven attribute means in linear testing

Classified Knowledge States:

Table 3.8 shows a high rate of matching in classified states by adaptive and linear modes.

Table 3.8: Comparison of frequencies classified in states by linear versus adaptive testing

		Adaptive Test State ID												Total
		1	2	3	4	6	7	8	10	13	15	17	30	
State ID	1	**87**	0	0	**2**	0	0	**1**	0	0	**1**	0	0	**91**
	2	0	**2**	0	0	0	0	0	0	0	0	0	0	**2**
Linear Test	3	0	0	**8**	0	0	0	0	0	0	0	0	0	**8**
	4	0	0	0	**0**	**2**	**1**	0	0	0	0	0	0	**3**
	6	0	0	0	0	**4**	0	0	**1**	0	0	0	0	**5**
	7	0	0	0	0	0	**2**	0	0	0	0	0	0	**2**
	8	**3**	0	0	0	0	0	**0**	0	0	0	0	0	**3**
	13	0	0	0	0	0	0	0	0	**1**	0	0	0	**1**
	17	0	0	0	0	0	0	0	0	0	0	**1**	0	**1**
	30	0	0	0	0	0	0	0	0	0	0	0	**2**	**2**
Total		90	2	8	2	6	3	1	1	1	1	1	2	

From the table above, 87 students are classified into KS1 by both of the two modes of testing. Only 3 students are classified by adaptive testing to KS1 while those students are classified to other states by linear testing. Similarly, only 4 students are classified to KS1 by linear testing, yet are not classified to KS1 by adaptive testing (2to KS4 and one each to KS8 and KS15).

Overall, only 10 students did not have matching classifications with respect to knowledge states out of the 119 students. In terms of attribute mastery, eight out of those 10 students had discrepancies in classification between adaptive and linear test results for only one attribute, while the other two did not match on determining mastery with respect to two attributes.

Table 3.8 shows how efficiently adaptive testing can conduct classifications as compared with linear testing, which requires 20 items. The average length of adaptive tests is 6.1 items, while linear testing requires 20 items. This is a dramatic reduction in testing without sacrifice of precision in classification.

Table 3.9: Number of items administered

Number of Items	Frequency	Percentage
5	84	70.6
6	5	4.2
7	2	1.7
8	7	5.9
9	1	0.8
10	19	16

4. RESULTS

4.1. PRESCRIPTION REPORTS FOR EACH CLASS

4.1.1PRESCRIPTION REPORT: CLASS A

The following fraction test was given to Class A, B, C, D and E, a total of 119 students. In Class A 23 students participated in the assessment. The items are as follows:

1) $\frac{5}{3}-\frac{3}{4}=?$	11) $4\frac{1}{3}-2\frac{4}{3}=?$
2) $\frac{3}{4}-\frac{3}{8}=?$	12) $\frac{11}{8}-\frac{1}{8}=?$
3) $\frac{5}{6}-\frac{1}{9}=?$	13) $3\frac{3}{8}-2\frac{5}{6}=?$
4) $4\frac{3}{5}-3\frac{4}{10}=?$	14) $3\frac{4}{5}-3\frac{2}{5}=?$
5) $3\frac{1}{2}-2\frac{3}{2}=?$	15) $2-\frac{1}{3}=?$
6) $\frac{6}{7}-\frac{4}{7}=?$	16) $4\frac{5}{7}-1\frac{4}{7}=?$
7) $3-2\frac{1}{5}=?$	17) $7\frac{3}{5}-\frac{4}{5}=?$
8) $\frac{3}{2}-\frac{3}{2}=?$	18) $4\frac{1}{10}-2\frac{8}{10}=?$
9) $3\frac{7}{8}-2=?$	19) $4-1\frac{4}{3}=?$
10) $4\frac{4}{12}-2\frac{7}{12}=?$	20) $4\frac{1}{3}-1\frac{5}{3}=?$

The test measured the following knowledge and skill components (attributes) :

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

The diagnostic testing model used to analyze the student's performance data. First we coded each item in terms of the attributes involved. Then looking at the pattern of correct and incorrect responses over 20 items we calculated each student's level of mastery on each of these attributes. These attributes mastery levels are measured by the probability of using them correctly when a student answers the items involving them, so their range is between 0 and 1. The cut of 0.80is used for determining whether a student mastered a given attribute or not.

Table 4.1: The attribute mastery levels for each individual student in Class A

Class A	A1	A2	A3	A4	A5	A6	A7
117	0.99	1	0.98	1	1	1	1
118	0.99	1	0.98	1	1	0.96	1
142	0.99	1	1	1	1	0.96	1
155	0.81	1	0.5	1	0	0.65	1
186	0.99	1	1	1	1	1	1
332	0.99	1	0.98	1	1	0.96	1
441	0.99	1	0.98	1	1	0.96	1
454	0.99	1	1	1	1	1	1
461	0.99	1	1	1	1	0.96	1
476	0.99	1	0.98	1	1	1	1
482	0.99	1	1	1	1	1	1
484	0.99	1	0.98	1	1	1	1
561	0.99	1	0.98	1	1	1	1
571	0.99	1	1	1	1	1	1
573	0.99	1	0.98	1	1	1	1
609	0.07	1	0.98	1	1	0.96	1
662	0.99	1	1	1	1	1	1
701	0.53	0.09	0.49	1	0.48	0.88	1
776	0.99	1	0.98	1	1	1	1
796	0.99	1	0.96	1	0.94	0.96	1
825	0.85	1	0.01	1	1	1	1
848	0.99	1	0.98	1	1	1	1
1709	0.24	1	0.14	1	0.75	0.94	1
1719	0.61	1	0.52	1	0.46	0.5	1

The color coding is determined as follows:

Green: Attribute mastery level is higher than .80. The student has mastered the attribute.

Yellow: Attribute mastery level is between .50 and .80. The student needs more practice to master the attribute

Red: Attribute mastery level is less than .50.The student has no mastered the attribute.

Table 4.2: The number of students in each of these three categories in Class A

Class A	A1	A2	A3	A4	A5	A6	A7
Mastered	20	23	19	24	20	22	24
Needs Practice	2	0	3	0	1	2	0
Not Mastered	2	1	2	0	3	0	0

Here we also provide the mean attribute mastery levels and total scores for our class along the same statistics for the total sample. Not that total score indicates the number of correct items.

On average, the students in Class A did well on the item fraction subtraction test. The means of attributes are above .88, above the cut off point .80 for mastery.

Table 4.3: Means and Standard Deviations (in parenthesis) of Total Score and Seven Attributes for Class A and Total Sample

	Total Score	A1	A2	A3	A4	A5	A6	A7
Class A	16.35 (3.01)	0.90 (.21)	0.96 (.19)	0.88 (.25)	1.00 (.00)	0.91 (.25)	0.94 (.12)	1.00 (.00)
Total	16.82 (3.5)	0.89 (.21)	0.98 (.13)	0.87 (.24)	0.98 (.11)	0.91 (.25)	0.93 (.18)	0.99 (.12)

The same descriptive information is displayed below as a graph:

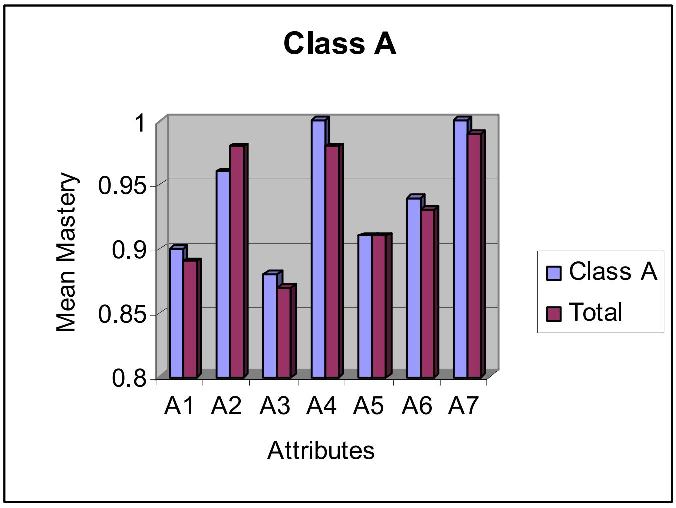


Figure 4.1 Descriptive information of Class A

Conclusion of Class A:

Class A did very well on all seven attributes on average. There was no statistically significant difference between the average performance of Class A and the average performance of other five classes on any of the seven attributes the entire class showed mastery on attributes A4 and A7. There are six students who need practice and / or instruction on one or more attributes. Below is a table that displays examples of items for practicing each attributes:

4.1.2 PRESCRIPTION RPORT: CLASS B

The following fraction test was given to Class A, B, C, D and E, a total of 119 students. In Class B 24 students participated in the assessment. The items are as follows:

1) $\frac{5}{3}-\frac{3}{4}=?$	11) $4\frac{1}{3}-2\frac{4}{3}=?$
2) $\frac{3}{4}-\frac{3}{8}=?$	12) $\frac{11}{8}-\frac{1}{8}=?$
3) $\frac{5}{6}-\frac{1}{9}=?$	13) $3\frac{3}{8}-2\frac{5}{6}=?$
4) $4\frac{3}{5}-3\frac{4}{10}=?$	14) $3\frac{4}{5}-3\frac{2}{5}=?$
5) $3\frac{1}{2}-2\frac{3}{2}=?$	15) $2-\frac{1}{3}=?$
6) $\frac{6}{7}-\frac{4}{7}=?$	16) $4\frac{5}{7}-1\frac{4}{7}=?$
7) $3-2\frac{1}{5}=?$	17) $7\frac{3}{5}-\frac{4}{5}=?$
8) $\frac{3}{2}-\frac{3}{2}=?$	18) $4\frac{1}{10}-2\frac{8}{10}=?$
9) $3\frac{7}{8}-2=?$	19) $4-1\frac{4}{3}=?$
10) $4\frac{4}{12}-2\frac{7}{12}=?$	20) $4\frac{1}{3}-1\frac{5}{3}=?$

The test measured the following knowledge and skill components (attributes) :

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

The diagnostic testing model used to analyze the student's performance data. First we coded each item in terms of the attributes involved. Then looking at the pattern of correct and incorrect responses over 20 items we calculated each student's level of mastery on each of these attributes. These attributes mastery levels are measured by the probability of using them correctly when a student answers the items involving them, so their range is between 0 and 1. The cut of 0.80is used for determining whether a student mastered a given attribute or not.

Table 4.4: The attribute mastery levels for each individual student inC lass B

Class B	A1	A2	A3	A4	A5	A6	A7
974	0.27	1	0.5	1	0.01	0.05	1
606	0.99	1	1	1	1	0.23	1
1609	0.97	1	0.96	1	0.96	0.25	1
196	0.56	0.52	0.49	0.57	0.48	0.41	0.22
315	0.5	0.5	0.5	0.5	0.5	0.5	0
1624	0.56	1	0.5	1	0	0.73	1
215	0.97	1	0.85	0.99	1	0.93	1
105	0.97	1	0.85	0.93	1	0.94	1
676	0.83	1	0.12	1	0.95	0.96	1
960	0.83	1	0.12	1	0.95	0.96	1
381	0.07	1	0.98	1	1	0.96	1
770	0.99	1	0.98	1	1	0.96	1
230	0.99	1	1	1	1	0.96	1
794	0.08	1	0.95	0	1	0.99	1
421	0.99	1	0.98	0.94	1	1	1
310	0.99	1	1	0.95	1	1	1
327	0.97	1	0.85	1	1	1	1
638	0.99	1	0.98	1	1	1	1
762	0.99	1	0.98	1	1	1	1
216	0.99	1	1	1	1	1	1
319	0.99	1	1	1	1	1	1
414	0.99	1	1	1	1	1	1
627	0.99	1	1	1	1	1	1
831	0.99	1	1	1	1	1	1

The color coding is determined as follows:

Green: Attribute mastery level is higher than .80. The student has mastered the attribute.

Yellow: Attribute mastery level is between .50 and .80. The student needs more practice to master the attribute

Red: Attribute mastery level is less than .50. The student has no mastered the attribute.

Table 4.5: The number of students in each of these three categories in Class B

Class B	A1	A2	A3	A4	A5	A6	A7
Mastered	19	23	19	22	21	19	23
Needs Practice	3	2	4	2	2	2	0
Not Mastered	3	0	2	1	2	4	2

Here we also provide the mean attribute mastery levels and total scores for our class along the same statistics for the total sample. Not that total score indicates the number of correct items.

On average, the students in Class B did well on the item fraction subtraction test. The means of attributes are above .81, above the cut off point .80 for mastery.

Table 4.6: Means and Standard Deviations (in parenthesis) of Total Score and Seven Attributes for Class B and Total Sample

	Total Score	A1	A2	A3	A4	A5	A6	A7
Class B	14.79 (4.83)	.81 (.30)	.96 (.14)	.82 (.28)	.92 (.23)	.87 (.30)	.83 (.30)	.93 (.25)
Total	16.82 (3.5)	0.89 (.21)	0.98 (.13)	0.87 (.24)	0.98 (.11)	0.91 (.25)	0.93 (.18)	0.99 (.12)

The same descriptive information is displayed below as a graph:

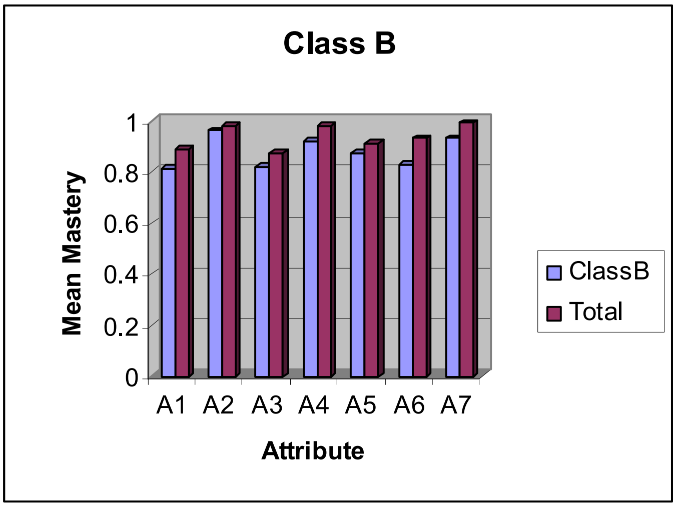


Figure 4.2: Descriptive information of Class B

Conclusion for Class B:

Class B did very well on all seven attributes on average. The mean attribute mastery probabilities for students in this class were lower compared to Class C and Class E only on A6. The students in this class performed especially well on A2. There are ten students who need practice and / or instruction on one or more attributes. Below is a table that displays examples of items for practicing each attributes:

4.1.3 PRESCRIPTION REPORT: CLASS C

The following fraction test was given to Class A, B, C, D and E, a total of 119 students. In Class C 24 students participated in the assessment. The items are as follows:

1) $\frac{5}{3}-\frac{3}{4}=?$	11) $4\frac{1}{3}-2\frac{4}{3}=?$
2) $\frac{3}{4}-\frac{3}{8}=?$	12) $\frac{11}{8}-\frac{1}{8}=?$
3) $\frac{5}{6}-\frac{1}{9}=?$	13) $3\frac{3}{8}-2\frac{5}{6}=?$
4) $4\frac{3}{5}-3\frac{4}{10}=?$	14) $3\frac{4}{5}-3\frac{2}{5}=?$
5) $3\frac{1}{2}-2\frac{3}{2}=?$	15) $2-\frac{1}{3}=?$
6) $\frac{6}{7}-\frac{4}{7}=?$	16) $4\frac{5}{7}-1\frac{4}{7}=?$
7) $3-2\frac{1}{5}=?$	17) $7\frac{3}{5}-\frac{4}{5}=?$
8) $\frac{3}{2}-\frac{3}{2}=?$	18) $4\frac{1}{10}-2\frac{8}{10}=?$
9) $3\frac{7}{8}-2=?$	19) $4-1\frac{4}{3}=?$
10) $4\frac{4}{12}-2\frac{7}{12}=?$	20) $4\frac{1}{3}-1\frac{5}{3}=?$

The test measured the following knowledge and skill components (attributes) :

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

The diagnostic testing model used to analyze the student's performance data. First we coded each item in terms of the attributes involved. Then looking at the pattern of correct and incorrect responses over 20 items we calculated each student's level of mastery on each of these attributes. These attributes mastery levels are measured by the probability of using them correctly when a student answers the items involving them, so their range is between 0 and 1. The cut of 0.80is used for determining whether a student mastered a given attribute or not.

Table 4.7: The attribute mastery levels for each individual student inClass C

Studen ID	A1	A2	A3	A4	A5	A6	A7
123	0.99	1	1	1	1	1	1
137	0.99	1	1	1	1	1	1
240	0.99	1	1	1	1	1	1
276	0.99	1	1	1	1	1	1
316	0.99	1	0.98	1	1	1	1
320	0.99	1	1	1	0	1	1
362	0.99	1	1	1	1	1	1
366	0.62	1	0.41	1	0.24	0.62	1
393	0.88	1	0.22	0.96	1	0.95	1
436	0.99	1	1	1	1	1	1
439	0.99	1	1	1	1	1	1
494	0.08	0.99	0.5	1	0.01	0.89	1
533	0.79	1	0.95	1	1	1	1
646	1	1	0.99	1	1	1	1
675	0.99	1	1	0.93	1	0.94	1
739	0.99	1	1	1	1	1	1
806	0.99	1	0.98	1	1	1	1
823	0.99	1	1	1	1	0.96	1
933	0.67	1	0.54	1	1	1	1
941	0.99	1	1	1	1	1	1
993	0.99	1	0.98	1	1	1	1
1616	0.99	1	0.23	1	1	1	1
1638	0.99	1	0.98	1	1	1	1
1732	0.99	1	1	1	1	1	1

The color coding is determined as follows:

Green: Attribute mastery level is higher than .80. The student has mastered the attribute.

Yellow: Attribute mastery level is between .50 and .80. The student needs more practice to master the attribute

Red: Attribute mastery level is less than .50. The student has no mastered the attribute.

Table 4.8: The number of students in each of these three categories in Class C

Class C	A1	A2	A3	A4	A5	A6	A7
Mastered	21	24	19	24	22	23	24
Needs Practice	2	0	2	0	0	1	0
Not Mastered	1	0	3	0	2	0	0

Here we also provide the mean attribute mastery levels and total scores for our class along the same statistics for the total sample. Not that total score indicates the number of correct items.

On average, the students in Class C did well on the item fraction subtraction test. The means of attributes are above .81, above the cut off point .80 for mastery.

Table 4.9: Means and Standard Deviations (in parenthesis) of Total Score and Seven Attributes for Class C and Total Sample

	Total Score	A1	A2	A3	A4	A5	A6	A7
Class C	17.04 (2.48)	.91 (.20)	1 (.00)	.86 (.28)	.99 (.2)	.92 (.25)	.97 (.08)	.1.0 (.00)
Total	16.82 (3.5)	0.89 (.21)	0.98 (.13)	0.87 (.24)	0.98 (.11)	0.91 (.25)	0.93 (.18)	0.99 (.12)

The same descriptive information is displayed below as a graph:

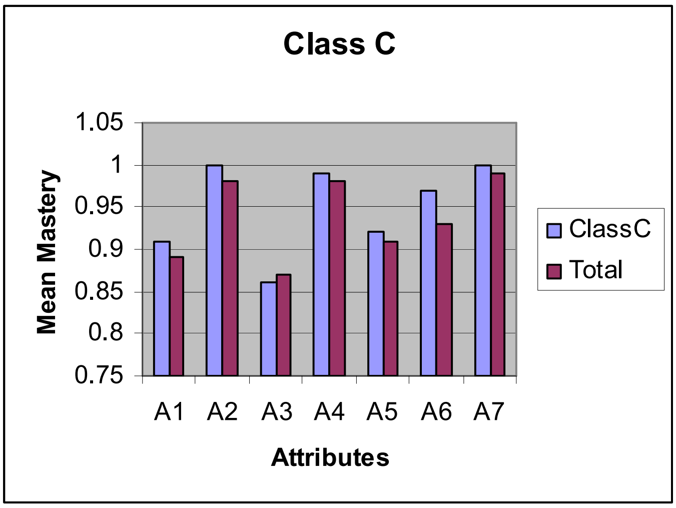


Figure 4.3: Descriptive information of Class C

Conclusion of Class C:

Class C did very well on all seven attributes on average. The mean attribute mastery probabilities for students in this class were lower compared to Class B only on A6 and at the same level with all other attributes. The entire class showed mastery on attributes A2, A4 and A7.There are five students who need practice and / or instruction on one or more attributes. Below is a table that displays examples of items for practicing each attributes:

4.1.4. PRESCRIPTION REPORT: CLASS D

The following fraction test was given to Class A, B, C, D and E, a total of 119 students. In Class D 24 students participated in the assessment. The items are as follows:

---

1) $\frac{5}{3}-\frac{3}{4}=?$

2) $\frac{3}{4}-\frac{3}{8}=?$

3) $\frac{5}{6}-\frac{1}{9}=?$

4) $4\frac{3}{5}-3\frac{4}{10}=?$

5) $3\frac{1}{2}-2\frac{3}{2}=?$

6) $\frac{6}{7}-\frac{4}{7}=?$

7) $3-2\frac{1}{5}=?$

8) $\frac{3}{2}-\frac{3}{2}=?$

9) $3\frac{7}{8}-2=?$

10) $4\frac{4}{12}-2\frac{7}{12}=?$

11) $4\frac{1}{3}-2\frac{4}{3}=?$

12) $\frac{11}{8}-\frac{1}{8}=?$

13) $3\frac{3}{8}-2\frac{5}{6}=?$

14) $3\frac{4}{5}-3\frac{2}{5}=?$

15) $2-\frac{1}{3}=?$

16) $4\frac{5}{7}-1\frac{4}{7}=?$

17) $7\frac{3}{5}-\frac{4}{5}=?$

18) $4\frac{1}{10}-2\frac{8}{10}=?$

19) $4-1\frac{4}{3}=?$

20) $4\frac{1}{3}-1\frac{5}{3}=?$

The test measured the following knowledge and skill components (attributes) :

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

The diagnostic testing model used to analyze the student's performance data. First we coded each item in terms of the attributes involved. Then looking at the pattern of correct and incorrect responses over 20 items we calculated each student's level of mastery on each of these attributes. These attributes mastery levels are measured by the probability of using them correctly when a student answers the items involving them, so their range is between 0 and 1. The cut of 0.80is used for determining whether a student mastered a given attribute or not.

Table 4.10:The attribute mastery levels for each individual student in our Class D

Class D	A1	A2	A3	A4	A5	A6	A7
104	0.99	1	0.98	1	1	0.96	1
108	0.88	1	0.21	1	1	1	1
109	0.52	1	0.96	1	1	0.96	1
111	0.99	1	1	1	1	1	1
221	0.99	1	0.98	1	1	1	1
234	0.99	1	1	1	1	1	1
262	0.99	1	0.98	1	1	1	1
286	0.99	1	1	1	1	1	1
372	0.99	1	0.98	1	1	1	1
373	0.99	1	1	1	1	1	1
376	0.99	1	1	1	1	0.23	1
431	0.99	1	1	1	1	1	1
435	0.99	1	1	1	1	1	1
520	0.87	1	1	1	1	1	1
530	0.99	1	1	1	1	1	1
557	0.99	1	0.98	1	1	1	1
581	0.85	1	0.38	1	0.28	0.53	1
655	0.56	1	0.5	1	0.01	0.73	1
663	0.99	1	1	1	1	0.96	1
668	0.99	1	1	1	1	1	1
741	0.99	1	1	1	1	1	1
837	0.99	1	0.98	1	1	1	1
869	0.99	1	1	1	1	1	1
877	0.99	1	1	1	1	1	1

The color coding is determined as follows:

Green: Attribute mastery level is higher than .80. The student has mastered the attribute.

Yellow: Attribute mastery level is between .50 and .80. The student needs more practice to master the attribute

Red: Attribute mastery level is less than .50. The student has no mastered the attribute.

Table 4.11: The number of students in each of these three categoriesin Class D

Class D	A1	A2	A3	A4	A5	A6	A7
Mastered	22	24	21	24	22	22	24
Needs Practice	2	0	1	0	0	2	0
Not Mastered	0	0	2	0	2	0	0

Here we also provide the mean attribute mastery levels and total scores for our class along the same statistics for the total sample. Not that total score indicates the number of correct items.

On average, the students in Class D did well on the item fraction subtraction test. The means of attributes are above .81, above the cut off point .80 for mastery.

Table 4.12: Means and Standard Deviations (in parenthesis) of Total Score and Seven Attributes for Class A and Total Sample

	Total Score	A1	A2	A3	A4	A5	A6	A7
Class D	17.92 (2.5)	.94 (.13)	1.00 (.00)	.91 (.26)	1.00 (.00)	.93 (.25)	.93 (.18)	1.00 (.00)
Total	16.82 (3.5)	0.89 (.21)	0.98 (.13)	0.87 (.24)	0.98 (.11)	0.91 (.25)	0.93 (.18)	0.99 (.12)

The same descriptive information is displayed below as a graph:

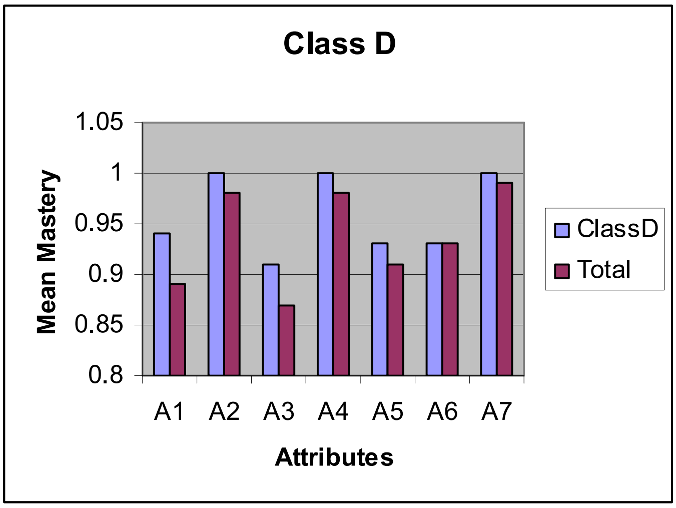


Figure 4.4: Descriptive information of Class D

Conclusion of Class D:

Class D did very well on all seven attributes on average. There was no statistically significant difference between the average performance of Class D and the average performance of other four classes on any of the seven attributes. The entire class showed mastery on attributes A2, A4 and A7. There are five students who need practice and / or instruction on one or more attributes.

4.1.5. PRESCRIPTION REPORT: CLASS E

The following fraction test was given to Class A, B, C, D and E, a total of 119 students. In Class E 24 students participated in the assessment. The items are as follows:

1) $\frac{5}{3} - \frac{3}{4} = ?$	11) $4\frac{1}{3} - 2\frac{4}{3} = ?$
2) $\frac{3}{4} - \frac{3}{8} = ?$	12) $\frac{11}{8} - \frac{1}{8} = ?$
3) $\frac{5}{6} - \frac{1}{9} = ?$	13) $3\frac{3}{8} - 2\frac{5}{6} = ?$
4) $4\frac{3}{5} - 3\frac{4}{10} = ?$	14) $3\frac{4}{5} - 3\frac{2}{5} = ?$
5) $3\frac{1}{2} - 2\frac{3}{2} = ?$	15) $2 - \frac{1}{3} = ?$
6) $\frac{6}{7} - \frac{4}{7} = ?$	16) $4\frac{5}{7} - 1\frac{4}{7} = ?$
7) $3 - 2\frac{1}{5} = ?$	17) $7\frac{3}{5} - \frac{4}{5} = ?$
8) $\frac{3}{2} - \frac{3}{2} = ?$	18) $4\frac{1}{10} - 2\frac{8}{10} = ?$
9) $3\frac{7}{8} - 2 = ?$	19) $4 - 1\frac{4}{3} = ?$
10) $4\frac{4}{12} - 2\frac{7}{12} = ?$	20) $4\frac{1}{3} - 1\frac{5}{3} = ?$

The test measured the following knowledge and skill components (attributes) :

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

The diagnostic testing model used to analyze the student's performance data. First we coded each item in terms of the attributes involved. Then looking at the pattern of correct and incorrect responses over 20 items we calculated each student's level of mastery on each of these attributes. These attributes mastery levels are measured by the probability of using them correctly when a student answers the items involving them, so their range is between 0 and 1. The cut of 0.80is used for determining whether a student mastered a given attribute or not.

Table 4.13: The attribute mastery levels for each individual student in our Class E

Class E	A1	A2	A3	A4	A5	A6	A7
156	0.99	1	1	1	1	1	1
209	0.99	1	1	1	1	1	1
224	0.79	1	0.95	1	1	1	1
235	0.99	1	1	1	1	1	1
265	0.99	1	1	1	1	1	1
323	0.99	1	1	0.95	1	1	1
364	0.99	1	1	1	1	1	1
367	0.99	1	1	1	0.99	0.96	1
426	0.99	1	1	1	1	1	1
452	0.99	1	1	1	1	0.96	1
472	0.99	1	1	1	1	1	1
507	0.99	1	1	1	1	1	1
546	0.99	1	1	1	1	1	1
608	0.99	1	1	1	1	1	1
690	0.99	1	1	1	1	1	1
754	0.99	1	1	1	1	1	1
801	0.74	0.76	0.5	1	0.12	0.72	1
926	0.99	1	1	1	1	1	1
930	0.87	1	0.46	1	1	1	1
936	0.99	1	1	1	1	1	1
969	1	1	0.99	1	1	1	1
976	0.56	0.21	0.5	1	0.4	0.86	1
1618	0.99	1	1	1	1	1	1
1709	0.24	0.98	0.14	1	0.75	0.94	1

The color coding is determined as follows:

Green: Attribute mastery level is higher than .80. The student has mastered the attribute.

Yellow: Attribute mastery level is between .50 and .80. The student needs more practice to master the attribute

Red: Attribute mastery level is less than .50. The student has no mastered the attribute.

Table 4.14: The number of students in each of these three categories in Class E

Class E	A1	A2	A3	A4	A5	A6	A7
Mastered	21	22	20	24	21	23	24
Needs Practice	2	1	3	0	1	1	0
Not Mastered	1	1	1	0	2	0	0

Here we also provide the mean attribute mastery levels and total scores for our class along the same statistics for the total sample. Not that total score indicates the number of correct items.

On average, the students in Class E did well on the item fraction subtraction test. The means of attributes are above .81, above the cut off point .80 for mastery.

Table 4.15:Means and Standard Deviations (in parenthesis) of Total Score and Seven Attributes for Class E and Total Sample

	Total Score	A1	A2	A3	A4	A5	A6	A7
Class E	17.75 (3.38)	.92 (.18)	.96 (.17)	.90 (.21)	1.00 (.00)	.93 (.22)	.98 (.06)	1.00 (.00)
Total	16.82 (3.5)	0.89 (.21)	0.98 (.13)	0.87 (.24)	0.98 (.11)	0.91 (.25)	0.93 (.18)	0.99 (.12)

The same descriptive information is displayed below as a graph:

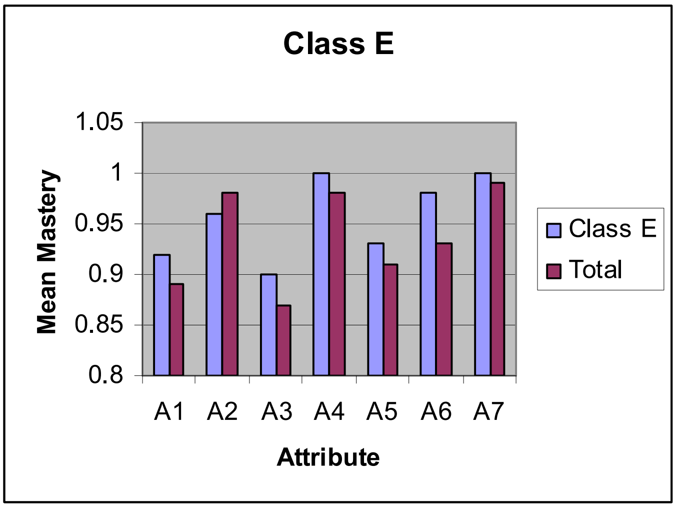


Figure 45 : Descriptive information of Class E

Conclusion of Class E:

Class E did very well on all seven attributes on average. The mean attribute mastery probabilities for students in this class were lower compared to Class E and the average performance of other four classes on any of the seven attributes. The entire class showed mastery on attributes A4 and A7. There are four students who need practice and / or instruction on one or more attributes. Below is a table that displays examples of items for practicing each attributes:

4.2.CLASS REPORTS FOR THE CLASSROOM T EACHER AT 7^{TH} GRADE

The results from the rule space model can be used for preparing a variety of reports that are tailored to different groups of test users. The purposes for using test reports may vary among different groups of test users.

The optimal use of results should be recommended. If the audience is higher educational institutes, test results are used for selection or placement of applicants. Individual examines in high schools may use test results for guiding themselves for further study or remediation , and teachers for evaluating their instructions, for designing of curricula and future instruction planning. The test results can also be used for preparing reports for group performance. Summary statistics of attribute-level performance as well as item level performance can be useful for schools. And for every student diagnostic scoring reports are prepared for the parents.

The subtraction of fraction test was given to 119 students at 7^{th} grade in Bahcesehir School. Two kinds of reports are given for the classes and the classroom teachers. For the 119 student's individual diagnostic scoring reports prepared, and the students can understand their levels. Also the students and parents can see the results from the school web page.

The numberof correct answers is minimum 7 and maximum 20, out of 20 questions. 38 students no mastered or need some practice to solve the items.

There are 20 items and 7 attributes such as in the subtraction of fractions, then the possible mastery patterns will be $2^7 = 128$.

4.2.1. THE CLASS A REPORT FOR THECLASSROOM TEACHER

Number of items: 20

Description of attributes:

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

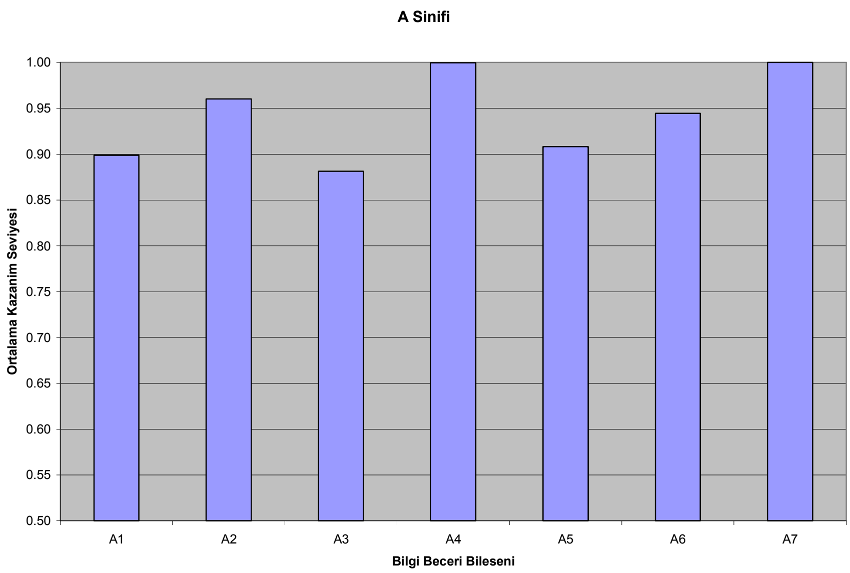


Figure 4.6: Class A teacher report

Skills that need remediation: (attribute probability < 0.50)

Other skills that could possibly use further practice: (0.50 < attribute prob. < 0.80)

Mastered skills : (attribute probability > 0.80)

4.2.2. THE CLASS B REPORT FOR THECLASSROOM TEACHER

Number of items: 20

Description of attributes:

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

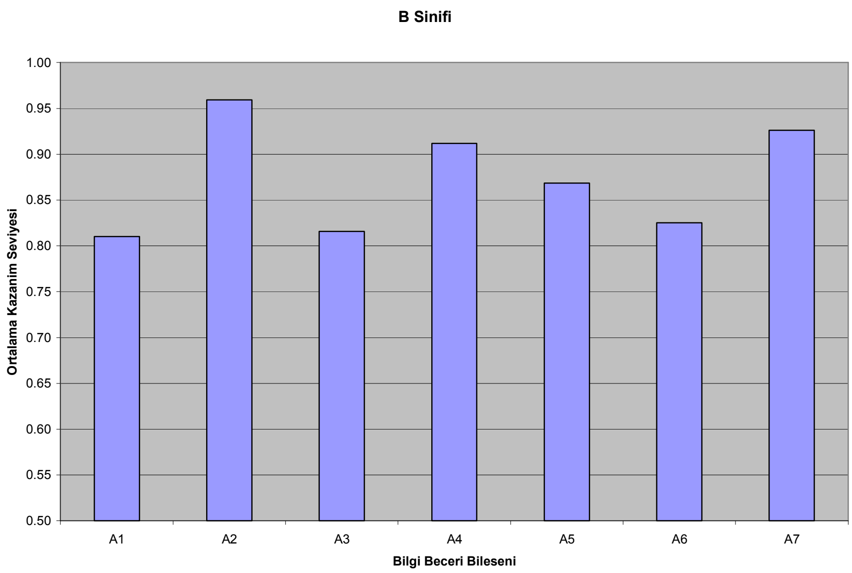


Figure 4.7: Class B teacher report

Skills that need remediation: (attribute probability < 0.50)

Other skills that could possibly use further practice: (0.50 < attribute prob. < 0.80)

Mastered skills: (attribute probability > 0.80)

4.2.3. THE CLASS C REPORT FOR THECLASSROOM TEACHER

Number of items: 20

Description of attributes:

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

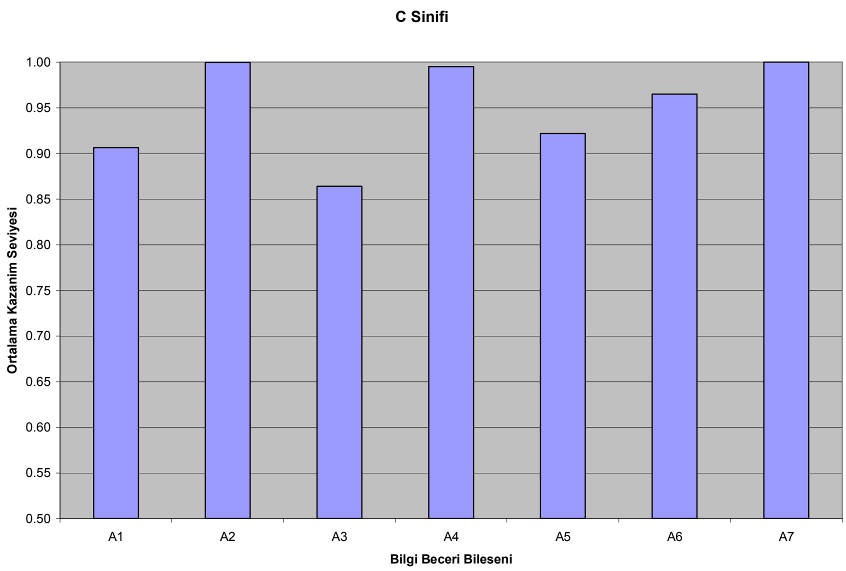


Figure 4.8: Class C teacher report

Skills that need remediation: (attribute probability < 0.50)

Other skills that could possibly use further practice: (0.50 < attribute prob. < 0.80)

Mastered skills : (attribute probability > 0.80)

4.2.4.THE CLASS D REPORT FOR THECLASSROOM TEACHER

Number of items: 20

Description of attributes:

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

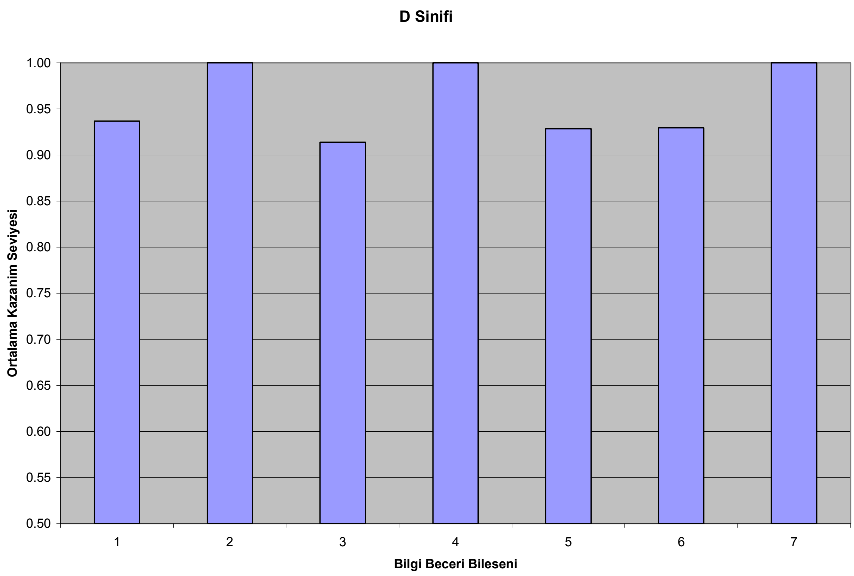


Figure 4.9: Class D teacher report

Skills that need remediation: (attribute probability < 0.50)

Other skills that could possibly use further practice: (0.50 < attribute prob. < 0.80)

Mastered skills : (attribute probability > 0.80)

4.2.5. THE CLASS E REPORT FOR THECL ASSROOM TEACHER

Number of items: 20

Description of attributes:

A 1: Convert a whole number to a fraction or a mixed number

A 2: Separate a whole number from a fraction

A 3: Simplify before subtraction

A4: Find a common denominator and make equivalent fractions

A5: Borrow 1 from a whole number part and add the denominator to the numerator

A6: Column borrow to subtract the second numerator from the first

A7: Subtracting the second numerator from the first

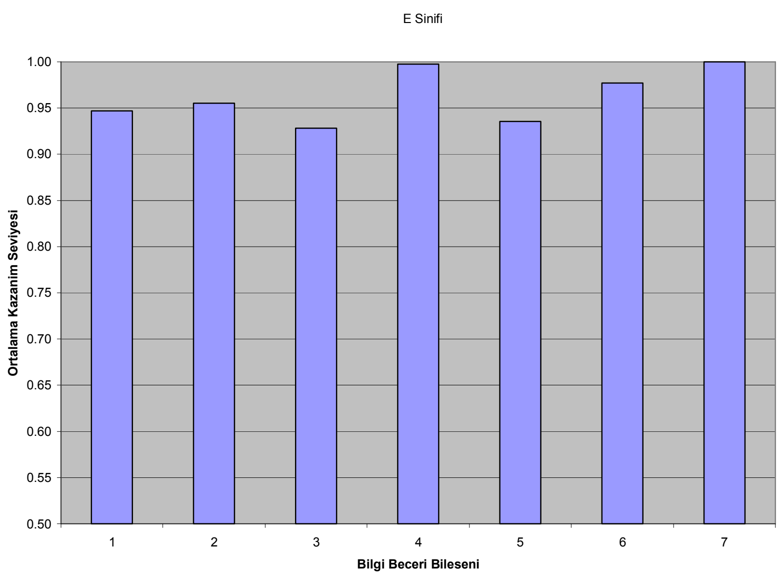


Figure 4.10: Class E teacher report

Skills that need remediation: (attribute probability < 0.50)

Other skills that could possibly use further practice: (0.50 < attribute prob. < 0.80)

Mastered skills: (attribute probability > 0.80)

4.3. DIAGNOSTIC SCORING REPORT FOR FRACTION SUBTRACTION SKILLS

The reports are prepared for the students who are interested in understanding their weaknesses and strengths. And personal targeted instruction is planned with this report.

Example of individual diagnostic report :

Student ID: 794 **Total Score:** 9 out of 20

Description ofattributes :

A 1: Convert a whole number to a fraction or a mixed number: **0.08**

A 2: Separate a whole number from a fraction: **1**

A 3: Simplify before subtraction: **0.95**

A4 : Find a common denominator and make equivalent fractions : **0**

A5: Borrow 1 from a whole number part and add the denominator to the numerator: **1**

A6: Column borrow to subtract the second numerator from the first: **0.99**

A7: Subtracting the second numerator from the first: **1**

Student ID: 581 **Total Score:** 13out of 20

Description ofattributes :

A 1: Convert a whole number to a fraction or a mixed number: **0.85**

A 2: Separate a whole number from a fraction: **1**

A 3: Simplify before subtraction: **0.38**

A4 : Find a common denominator and make equivalent fractions : **1**

A5: Borrow 1 from a whole number part and add the denominator to the numerator: **0.28**

A6: Column borrow to subtract the second numerator from the first: **0.53**

A7: Subtracting the second numerator from the first: **1**

These two cases illustrate the advantage of RSM over traditional methods. As seen above Student ID= 794 had 4 less correct items than Student ID= 581 although s / he mastered 1 more attribute than Student ID= 581 (the cut of 0.80 is used for mastery).This example shows how important diagnostic information is as feedback for instruction.

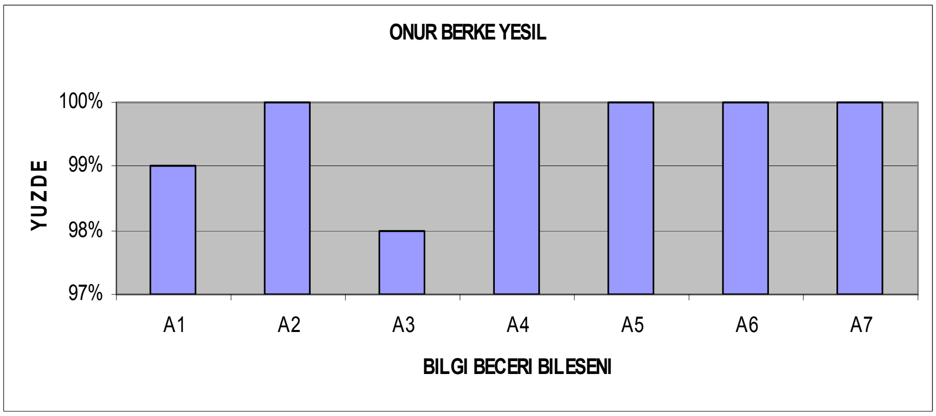
BAHCESEHIR KOLEJI

KESIRLER TESTI TANISAL OGRENCI RAPORU

Ogrenci Adı /Soyadı: **ONUR BERKE YESIL**
Okul Numarası / Sınıfı: **117 / 7-A**
Toplam dogru cevap sayısı: 18

A1: Bir tamsayıyı kesir veya bilesik kesir seklinde yazma: **.99**
A2: Bir kesrin icindeki tamı (butunu) ayırma: **1**
A3: Cıkarmadan once sadelestirme yapma: **.98**
A4: Ortak payda bulma: **1**
A5: Bilesik kesrin tam kısmından bir tam alma: **1**
A6: Cıkarmada tamı kesre katarak 1. paydan 2. payı cıkarma: **1**
A7: Basit kesirlerde cıkarma: **1**

Ortalama Bilgi Beceri Bilesenleri Kazanım Seviyesi



Ogrencinin Bilgi Beceri Bilesenleri Kazanım Seviyeleri Nasıl Okunmalı:

Eger seviye .8'den buyuk ise ------> Ogrenci Bilgi Beceri Bilesenini kazanmıs

Eger seviye .5 ile .8 arasında ise ---> Ogrencinin Bilgi Beceri Bilesenini kazanması icin pratik yapması gerekli

Eger seviye .5'den kucuk ise <.5 ---> Ogrenci Bilgi Beceri Bilesenini kazanamamıs

SONUC: Ogrencimiz Onur Berke Yesil'in almıs oldugu kesirler testi sonucunda, olculen butun Bilgi Beceri Bilesenine ait kazanımları tamdır.

Figure 4.11: Example of individual diagnostic report

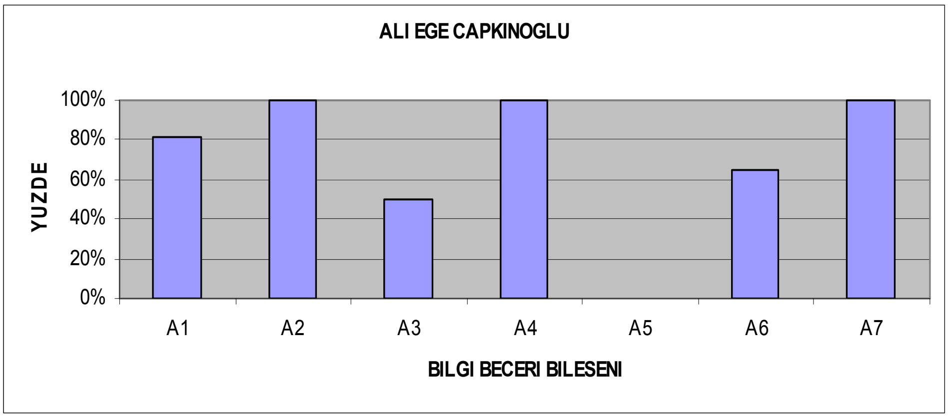
BAHCESEHIR KOLEJI

KESIRLER TESTI TANISAL OGRENCI RAPORU

Ogrenci Adı /Soyadı: **ALI EGE CAPKINOGLU**
Okul Numarası / Sınıfı: **155 / 7-A**
Toplam dogru cevap sayısı: 9

A1: Bir tamsayıyı kesir veya bilesik kesir seklinde yazma: **.81**
A2: Bir kesrin icindeki tamı (butunu) ayırma: **1**
A3: Cıkarmadan once sadelestirme yapma: **.50**
A4: Ortak payda bulma: **1**
A5: Bilesik kesrin tam kısmından bir tam alma: **.00**
A6: Cıkarmada tamı kesre katarak 1. paydan 2. payı cıkarma: **0.65**
A7: Basit kesirlerde cıkarma: **1**

Ortalama Bilgi Beceri Bilesenleri Kazanım Seviyesi



Ogrencinin Bilgi Beceri Bilesenleri Kazanım Seviyeleri Nasıl Okunmalı:

Eger seviye .8'den buyuk ise ------> Ogrenci Bilgi Beceri Bilesenini kazanmıs

Eger seviye .5 ile .8 arasında ise ---> Ogrencinin Bilgi Beceri Bilesenini kazanması icin pratik yapması gerekli

Eger seviye .5'den kucuk ise <.5 ---> Ogrenci Bilgi Beceri Bilesenini kazanamamıs

SONUC: Ogrencimiz Ali Ege Capkınoglu'nun almıs oldugu kesirler testi sonucunda, olculen Bilgi Beceri Bilesenine ait A3 ve A6 becerilerini kazanabilmesi icin pratik yapması gerekmektedir. Bilgi Beceri bilesenine ait A5 becerisini ise kazanamadıgı tespit edilmistir.

Figure 4.12: Example of individual diagnostic report

5. DISCUSSION

Bahcesehir Ilkogretim Okulu mathematics department teachers took three days seminar from Kumi Tatsuoka, learned using statistics, probability and computer coding with the Rule-Space Model which can determine individual strengths and weaknesses based on how a test taker responds to questions on the subtraction of fractions. Rule-Space Model (RSM) is seen as one of the most viable alternatives to the traditional one-dimensional models of Item Response Theory (IRT).

Sometimes students change their rules for choosing a certain answer for no reason. In addition, that student may not remember why he or she chose that particular answer. It's difficult to figure out the logic behind an answer choice because it's not something that they can observe while it's happening.

The methodology for the Rule-Space Model follows pattern analysis, which helps computers to read letters like humans do, each question has "feature variables." For example, when a computer distinguishes between the letter "A" and the letter "H," it looks at what differentiates the two characters. Those differences are the letters "features," or "attributes." These features are translated into code, a series of patterns of zeros and ones. Then, they match stored patterns with the new patterns. The computer processes them and reads the letters. It is the same approach used to read DNA.

Instead of looking at DNA or the alphabet, she is looking at questions such as, "Why can't Student A read?" or "Why can Student B do fraction problems?" To figure this out, she uses probability and statistics to determine the student's "knowledge state," which is a combination of the attributes that she compiles using pattern analysis?

With this system of scoring, each student gets a personal attribute mastery report showing how he or she did individually. This individual report breaks down the total score into categories of mastery and where the student needs improvement.

Students may have similar scores, but entirely different strengths and weaknesses; The Rule-Space Model analyzes the score and looks at the test takers' specific skills based on the answer choice he or she made. Two students who got 14 out of 20 on the subtraction of fractions have very different individual reports. For example, Student A, who scored 14, and he doesn't understand two attributes: simplify before subtracting and converting a whole number to a fraction or a mixed number. While student B also scored 14, he is not good borrowing 1 from a whole number part and column borrow to subtract the 2nd numerator from the first. This method tells schools and parents to see exactly what the lump score of 14 means. It tells them where the students went wrong and provides suggestions for improvement in specific areas. They can provide a snapshot of a student's abilities and provide suggestions to build on them.

How many possibilities are there for Student A who scored 14 out of 20?

$$C\begin{pmatrix} 20 \\ 14 \end{pmatrix} = \frac{20!}{14!.6!} = \frac{20.19.18.17.16!}{14!.6.5.4.3.2.1} = 38.760 \text{ Possibilities} \qquad (5.1)$$

Using RSM we can find easily which attributes are mastered or not. And RSM can give a diagnostic report for each possibility. RSM gave to the teacher a feedback about the class and the students.

The results from the rule space model can be used for preparing a variety of reports that are tailored to different groups of test users. The purposes for using the test reports may vary among different groups of test users. The optimal use of test results should be recommended. If the audience is higher educational institutes, test results are used for selection or placement of applicants. Individual examines in high schools may use test results for guiding themselves for further study or remediation, and teachers for evaluating their instructions, for designing of curricula and future instruction planning.

The RSM is a symbolic parametric model in which the performances on unobservable cognitive tasks are inferred from observable item scores.

REFERENCES

Books:

Brenhaum, M. K.,(1992) *Feedback from tests-A practical handbook for analyzing test results.*Tel-aviv

Carpenter,T.P. & Moser, J.M.,(1982) *The development of addition and subtraction problem solving skills.* NJ: Erlbaum

Corter, J. F. & Tatsuoka, K. K.,(2002) *Diagnostic Assessments for Mathematics Tests Grades 6-12* . Technical Report , College Board.

Duda, R.D. & Hart,P.E.,(1973) *Pattern classification and scene analysis.*Newyork

Fukunaga,K.,(1990) *Introduction to statistical pattern recognititon.(2nd edition).*Boston

Hogg,R.V.&Craig,A.T.,(1978) *Introduction to mathematical statistics.*Newyork, Macmillan

Sato, T., (1990) *An introduction to educational information technology*.Kawasaki,Japan

Tatsuoka,K.K.,(1990) *Toward an integration of item response theory andcagnitive analysis.* In N.Frederiksen,R.Glaser,A.Lesgold&M.C.Shafto (Eds) , *Diagnostic monitoring of skill and knowledge acquisition* pp.543–588. Hillsdale NJ: Erlbaum.

Tatsuoka, K. K.,(1991) *Boolean Algebra Appliedto Determination Of Universal Set Of Knowledge States*. Princeton, Nj: Educational Testing Service

Tatsuoka K. K., (1993) *Item construction and psychometric models appropriate for constructed responses.*Hillsdale ,NJ.

Tatsuoka K.K.,(In Press) *Statistical Pattern Recognition And Classification Of Latent Knowledge States. Cagnitively Diagnostic Assessment.* Mahwah, Nj: Erlbaum.

Periodicals :

Brenhaum, M. K.& Tatsuoka, K. K.,(in press) *Applying the rule space model to diagnose students , knowledge states in exponents.* Journal of Teaching Mathematics.

Hambleton, R.K., Jones, R.W., (1993 *Comparison of classical test theory and item response theory and their applications to test development*. Educational Measurement: Issues and Practice.12(3), pp. 38-47.

Tatsuoka, K. K., (1983) *Rule Space: An approach for dealing with misconceptions based on item response theory.* Journal of Educational Measurement, 20, pp. 34-38

Tatsuoka,K.K. & Linn,R.L.,(1983) *Indices for dedecting unusual patterns.* Applied Psychological Measurement ,7(1) ,pp. 81-96

Tatsuoka, K. K.,(1985) *A Probabilistic model for diagnosing misconceptions in the pattern classification approach.* Journal of Educational Statistics, 12(1), pp. 55-73

Tatsuoka,M.M. & Tatsuoka, K. K.,(1989) *Rule Space.* In S. Kotz & N.L.Johnson (Eds) Encyclopedia Of Statistical Sciences.(Vol.8, pp. 217-220). Newyork :Wiley.

Tatsuoka C., (2002) *Data analytic methods for latent partially ordered classification models:Applied statistics.* Journal of the Royal Statistical Society Series C,51, pp.337-350

Van Lehn,K.,(1982) *Bugs are not enough : Emprical studies of bugs, impasses and repairs in produceral skills.* Journal of Mathematical Behaviour, 3, pp. 3-71

Other Sources :

Gitomer, D.H.,(1988). *Sources of difficulty in calculus performance*.A paper presented at AREA in Neworleans

http://en.wikipedia.org/wiki/Classical_test_theory

http://en.wikipedia.org/wiki/Item_response_theory

http://ltj.sagepub.com/cgi/content/abstract/15/2/119

CURRICULUM VITAE

Name Surname	**:** H. Gulay TEZCAN
Permanent Adress	**:** Funda 0101D14 Kat:7 Daire:30 Bahcesehir / Buyukcekmece Istanbul
Birth Place and Date	**:** Eskisehir – 18.09.1969
	: Ingilizce
Primary School	**:** M. Aziz Bolel Ilkokulu; Eskisehir Osmangazi Ortaokulu; Eskisehir
Secondary School	**:** Ahmet Kanatlı Lisesi; Eskisehir
University	**:** ORTA DOGU TEKNIK UNIVERSITESI ; ANKARA / Matematik Ogretmenligi , 1986 – 1991

Working Life :

Ozel Ata Koleji	1991 – 1994
Istek Vakfı, Tarabya Kemal Ataturk Lisesi	1994 –1995
FMV Ayazaga Isık Lisesi	1995 – 1997
Ozel Bahcesehir Koleji	1997 – 2008
Ilkogretim Matematik Bolum Baskanlıgı	2006 – 2008
TED Istanbul	2008 – ......