HAZAL ELİF KARA

2016

STUDENTS' COGNITIVE LEVELS IN SCIENCE SUBTEST OF UNDERGRADUATE PLACEMENT EXAMINATION IN TURKEY

A MASTER'S THESIS

BY

HAZAL ELİF KARA

THE PROGRAM OF CURRICULUM AND INSTRUCTION
İHSAN DOĞRAMACI BILKENT UNIVERSITY
ANKARA

SEPTEMBER 2016

To my family…

STUDENTS' COGNITIVE LEVELS IN SCIENCE SUBTEST OF UNDERGRADUATE PLACEMENT EXAMINATION IN TURKEY

The Graduate School of Education

of

İhsan Doğramacı Bilkent University

by

Hazal Elif Kara

In Partial Fulfilment of the Requirements for the Degree of

Master of Arts

in

The Program of Curriculum and Instruction
İhsan Doğramacı Bilkent University
Ankara

September 2016

İHSAN DOĞRAMACI BILKENT UNIVERSITY
GRADUATE SCHOOL OF EDUCATION
STUDENTS' COGNITIVE LEVELS IN SCIENCE SUBTEST OF
UNDERGRADUATE PLACEMENT EXAMINATION IN TURKEY

Hazal Elif Kara
September 2016

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Curriculum and Instruction.

Asst. Prof. Dr. İlker Kalender

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Curriculum and Instruction.

Asst. Prof. Dr. Armağan Ateşkan

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Curriculum and Instruction.

Asst. Prof. Dr. Ela Ayşe Köksal

Approval of the Graduate School of Education

Director: Prof. Dr. Margaret K. Sands

ABSTRACT

STUDENTS' COGNITIVE LEVELS IN SCIENCE SUBTEST OF UNDERGRADUATE PLACEMENT EXAMINATION IN TURKEY

Hazal Elif Kara

M.A., Program of Curriculum and Instruction

Supervisor: Asst. Prof. Dr. İlker Kalender

September 2016

The Undergraduate Placement Examination (UPE) determines students' future careers in three hours. The content of the UPE, its quality, and the cognitive level of the items, must have high selectivity features in terms of getting students into higher education programs. The aim of this study is: to classify UPE science items in accordance with Bloom's Taxonomy, to determine the difficulty parameters of those items, and the cognitive levels of the students who take the UPE according to science subtests. Four science pre-service teachers classified 30 science items in regards to Bloom's Taxonomy. Quantitative exploratory research was employed as a research design. Quantitative analyses were carried out by using SPSS. Sample science responses of 3382 randomly selected students were analysed. It was found that there were 4 knowledge level items, 10 comprehension level items, 11 application level items, and 5 analysis level items as defined by Bloom's Taxonomy. Moreover, it was observed that students generally encountered difficulty when responding to science items. The results show that students were more likely to answer knowledge and comprehension level items.

Key words: Bloom's taxonomy, undergraduate placement examination, item analysis

ÖZET

ÜNİVERSİTE YERLEŞTİRME SINAVINA GİREN ÖĞRENCİLERİN, FEN ALANINDAKİ BİLİŞSEL SEVİYELERİ

Hazal Elif Kara

Yüksek Lisans, Eğitim Programları ve Öğretim

Tez Yöneticisi: Yrd. Doç. Dr. İlker Kalender

Eylül 2016

Üniversite yerleştirme sınavı, öğrencilerin tüm hayatları boyunca hangi mesleği icra edeceklerini yaklaşık olarak üç saat içinde şekillendiren bir sınavdır. Üniversite yerleştirme sınavının içeriği, soruların kalitesi ve ölçtüğü bilişsel seviyeler, üniversitelerin yüksek bilişsel seviyelerde öğrenci almaları açısından seçiciliği yüksek özellikte olmalıdır. Bu çalışmada üniversite yerleştirme sınavında sorulan fen sorularının Bloom Taksonomisine göre sınıflandırılması, soruların zorluk parametrelerinin bulunması, üniversite sınavına giren öğrencilerin bilişsel seviyelerinin bulunması. Sorular toplam dört stajyer biyoloji öğretmeni tarafından Bloom Taksonomisine göre incelenmiştir. Araştırma nicel keşfedici araştırma olarak tasarlanmıştır. Çalışmadaki nicel analizler SPSS paket programı kullanılarak yapılmıştır. Örneklem olarak rastgele seçilen 3382 öğrencinin fen cevapları incelenmiştir. Soruların 4 tanesinin bilgi, 10 tanesinin kavrama, 11 tanesinin uygulama ve 5 tanesinin analiz seviyesinde sorular olduğu bulunmuştur. Sorular çoğunlukla öğrenciler tarafından zor bulunmuştur. Sonuçlar göstermektedir ki, öğrenciler bilgi seviyesinde ve kavrama seviyesinde soruları yapmaya daha meyillidir.

Anahtar Kelimeler: Bloom taksonomisi, üniversiteye giriş sınavı, madde analizi

ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to Prof. Dr. Ali Doğramacı and Prof. Dr. M. K. Sands, and to the staff of İhsan Doğramacı Bilkent University Graduate School of Education for their help and support.

I would like to offer my sincerest thanks to my supervisor, Asst. Prof. Dr. İlker Kalender, for his efforts in assisting me with patience and devotion of his time throughout the process of writing this thesis. I would also like to offer my sincere thanks to the members of my committee, Asst. Prof. Dr. Armağan Ateşkan and Asst. Prof. Dr. Ela Ayşe Köksal for all their comments and feedback.

The most heartfelt thanks are for my wonderful family, my lovely mother, MEVLÜDE EVCAN, and my father, HIDIR KARA. I would like to thank my sibling EMRE KARA who always believed in and inspired me. I am heartily thankful to my cousin Dr. FÜRUZAN ÖZBEK LANDERS and her husband CHAD LANDERS for the endless support they had given me.

My many thanks to my lovely friends Özge Keşaplı Can, Bahar Kumandaş and Müjde Peder for their support and joy throughout this program.

Finally, I would like to offer my deepest gratitude to my husband, OĞULCAN CAF. I am eternally grateful for all of his suggestions and support. None of this would have been possible without his love and patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Introduction

Education is the means of any country who desires to reach the technological and scientific trends that shape the modern world. Symbiotically, these same trends are also beginning to significantly influence and change the current balance of education in the world today. Parallel to these developments, the importance of education has increased more than ever (Loveless & Ellis, 2002). Keeping up with these trends requires people who have the ability to think critically, investigate, inquiry, engage with existing knowledge in addition to new information, and to effectually utilize these skills in order to solve problems. Traditional education systems, wherein teachers lecture to multitudes of students in a classroom without the use of educational material and apply only standardized tests or written exams for assessments are not able to adapt to the developments in science and technology. This insufficiency in the face of contemporary developments hinders contributions that would further the development of the country. The countries, which give attention to this situation, have started to question and make innovations within their educational systems (Güneş, 2012). Unusual educational methods, which utilize the technological developments of the 21st century via computer animations, simulations, robotics, and technological laboratories, have been adopted in order to facilitate the learning and teaching process.

Learning is the process in which behavioural changes involving one's knowledge, skills, attitudes, and values, are generated through interactions with an environment (Hesapçıoğlu, 1994). The very basis of learning is therefore constituted of experience. The most intense and permanent part of learning takes place under the discipline of formal educational environments, although learning is a lifelong process cumulated throughout all encountered environments. Education begins in nursery school and is ongoing until one graduates from college. Having compulsory education up to a certain level allows more individuals to receive formal, disciplined, education within the community.

According to the Turkish Ministry of Education (2012), compulsory education is 11-12 years or more on average throughout the Post-Industrial world. Comparatively, the average years of schooling obtained by the adult population in Turkey were around 6-6.1 years prior to 2012, capturing only a virtual half of the average achieved by much of the developed world. Since the adoption of 12-year compulsory education in 2012, the average years of education obtained by the general population have increased while regional differences in enrolment rates have reduced (MoNE, 2012). When we look Industrial and Post-Industrial countries, almost all of them have set targets to increase significant portions of the population's completion of high school or university programs by increasing the average years of education and have taken measures to accomplish this. Despite the scope of compulsory education in primary and secondary levels, higher education has to have its central mission be not only the training of students in regards to vocation and the enrichment of individuals, but also with respects to comprehension of consistently renewed scientific knowledge, and to contribute to these developments through the utilization of technology. Higher education institutions evaluate, prompt, form and enrich scientific knowledge in light of the social dynamic (Çetinsaya, 2014).

Universities provide students an environment in which access to technology and scientific laboratories equip students with the background needed to approach their research and furthermore, their future vocations. According to the Turkish Council of Higher Education (2007), since the last quarter of the twentieth century, the transition to information-based societies among developed nations gave rise to a new global economic structure referred to as the knowledge economy. Associated with this new structure, the economic power of individuals and moreover, the competitiveness of the country itself is measured via knowledge and collective education levels, fuelled by human and social capital. This process has increased expectations of universities that have the primary responsibility of the production and sharing of knowledge. Higher education has become the central focus of societies the world over. This increase of expectations thrust upon universities can manifest in the following ways:

- Provide education to more people and wider age groups (massification)
- Expand educational programmes to include all of the new fields (academic expansion)
- Training graduates to find jobs and to turn research into knowledge and practice (relevance)
- Contribute to regional and national development by building strong bridges within the community
- Develop an open and transparent governance model accountable to its stakeholders (accountability)
- Meet all these expectations amid diminishing public resources (Turkish Council of Higher Education, 2007, p13)

Moreover, the 21st century has ushered in vast social, scientific, technological, environmental and economic transformations that require individuals to possess various skill sets in order to understand and adapt to such changes. Science education especially contributes to the improvements required to garner such skills and serves as one of the main objectives of education. These skills are primarily: analytical, critical thinking, creativity, innovation, problem solving, informatics, teamwork, entrepreneurship and responsibility (Head council of education and morality, 2013)

Under these growing expectations, the admission into a university is of considerable importance for students' professional and academic careers in Turkey, as well as all over the world. As of September 2016, there are 178 universities of which 69 are private and 109 are state, with roughly 6 million university students in attendance (Turkish Council of Higher Education, 2016).

While universities are providing a wide range of opportunities, they expect students to be able to do numerical operations, be literate, have fundamental knowledge in math and science, in addition to possessing information and communication technology skills. Consequently, those who are able to higher education must exhibit some distinct qualifications. As results of this, universities have maintained certain criteria by which to accept students into their educational programs. Due to the need for selection among applicants for potential students, various admission systems have been developed around the world.

It may be helpful to consider the university admission systems of different countries. In order to give some examples of worldwide university admission systems, some countries were chosen whose students exhibited high success rates in PISA (Program for International Student Assessment) (OECD, 2010).

In the Netherlands, in order to gain access to higher education through academic training colleges (VO), high school graduation certificates (VWO diploma) are required. To obtain this certificate, student need to go to an academic high school. Students attending vocational higher education (HBO) can transfer to universities (VO) after the first class. Because there is too much demand in medicine, students are subjected to an evaluation. The only requirement is to be 18 years old and be registered with the university (NUFFIC, 2016).

In Finland, there is a university entrance system. Students, who have successfully passed the high school completion (maturity) exam, in which candidates are required to take at least four subjects, then become eligible for admission. Matriculation examination is held twice a year and is done at the same time in schools. This is however, not the only selective process and universities are at liberty to compose their own entrance exams (Finland Ministry of Education, n.d).

The United States of America does not have a central admission system for higher education. Universities are free to set and apply their own admission requirements. Students are expected to have these requirements: the result of Scholastic Aptitude Test (SAT), which is a skill test or American College Testing (ACT), which is an achievement test, a reference letter, a personal statement showing an interest in, and academic potential for a specific area. (Burton & Ramist, 2001).

In England, students must successfully complete at least two courses of advanced-level (A-level) or an equivalent high school graduation exam for admission into the program. Universities usually set the requirements themselves. Advanced-level exams for university entrance is the decisive factor (Cambridge International, 2016). In addition, work experience, statement of purpose and a reference letter from a teacher would be effective. Universities may request further exams or documents, for example, the UK Clinical Aptitude Test (UKCAT) is requested for medical or dental schools. The processing of applications for universities is conducted by an organization called UCAS (University and Colleges Admissions Service). Universities are the authority to make the final decision on students' admissions. (UKCAT, 2016)

In Austria, students, who are successful in Matriculation examination, (Certificate of Reifeprüfung, Matura) usually, can register directly to the university. Matriculation examination includes written exams and oral exam. In Germany, high school graduates, who can successfully pass the final high school examination, have the right to take the certificate of Abitur. With this certificate, high school graduates can then gain access to college. Some departments, like medicine, are highly demanding. Students need to achieve a certain minimum score on the Abitur in order to be accepted into college (Gruber & Zdrahal-Urbanek, 2006).

In Turkey, Measurement, Selection, and Placement Center (MSPC) is the responsible institution for the university admissions system. The vision of the MSPC is to measure the knowledge and skills of individuals regarding the scientific method, contribute to the growth of competitive individuals, provide a positive contribution to the educational system and to develop an institution whose examination results are recognized at the international level (MSPC, 2012).

In Turkey, the university admission system has been revised throughout the years since its first inception in 1974 (Kutlu, 2003). There was only one centralized examination between 1975-1980 and 1999-2009. Between these two periods and beginning in 2010, the university admission system adopted two-stage examinations (MSPC, 2016).

In 2016, the Higher Education Transition Exam (HETE) is the first stage of the admission system for higher education and constitutes 40% of the placement scores. To pass the first stage, students must have a score of 180 or above. If students pass the HETE examination, they are permitted to take the second stage Undergraduate Placement Examination (UPE). HETE and UPE both have an effect on students' placement scores. However, there are some departments at universities which admit students into their four and two-year programs with only the HETE score. Both stages consist of multiple-choice questions from each of the following fields: mathematics, science, social science and Turkish language. The HETE includes items for screening. HETE has two sections: quantitative and verbal, which focus on students' ability to use basic concepts and the principles learned during the formal educational years. Mathematics and science subtests assess students' basic mathematical comprehension and reasoning ability to enforce scientific concepts through the use of generalized rules. Items in social sciences and Turkish subtests require students to make judgments using social studies concepts and generalizations in Turkish proficiency (MSPC, 2010). Additionally, items in the UPE are subject area-oriented and assesses students' achievements. All items are mainly assessing the cognitive domain.

In Turkey there is a lot of demand for higher education, whereas the lack of sufficient quotas to meet this demand requires the student selection process for universities to become increasingly exclusionary. The increasing demand by students and parents for tutoring with the aim of preparing for the UPE creates a negative impact on education in schools. Currently, the admission system has become inoperable in many ways. Berberoğlu (2012) stated these problems as:

- Exams only once a year
- Weak psychometric properties of tests
- The multiple-choice test format
- Non-standardized tests
- No information is provided on student's competencies
- Not possible to monitor the improvement of students
- Students who must retake the exam although they have already registered into a program
- Only academic-oriented assessments
- Unethical practises

However, the fact that there is a huge imbalance between the number of applicants and the available quotas in higher education institutions is probably one of the most important problems faced within the admission system. For example, in 2016, 2,256,377 students applied for the first stage, only 961,864 of which were placed into universities (MSPC, 2016). Similarly, the number of students, who took the HETE and were entitled to take the UPE, was 1,368,941 in 2015. The numbers of available quotas of universities were 648,781 students for state university and 153,965 students for private university. While the total quotas for universities were 802,746 and 773,176 students were placed in universities, however, 595,765 students were not admitted in any university at all (MSPC, 2015). Thus, due to its high-stakes nature, the HETE and the UPE questions hold a vital position in its role to assess students.

Since there is a huge imbalance between the number of students and the number of available quotas, the selection of students into higher education programs is more difficult than in many other developed countries. In such a competitive admission system, selecting the most appropriate students becomes a very important problem. Students should be selected based on the desired criteria designed by the system.

Since the university admission examinations are given only once a year, significant pressure is placed upon students. If the questions are not developed to accommodate the higher-order thinking levels of students, then the results should not be expected to select students with higher-order thinking skills. This study focuses on an analysis of the UPE questions in accordance with Bloom's taxonomy and the classification of students who take the same year exam according to their cognitive levels.

Background

Assessment is a term that is used to determine the level of student achievement in education (Leahy, Lyon, Thompson & Wiliam, 2005). It is a process of determining the level of student learning, skills and abilities by teachers or other professionals. Student achievement is measured by using formal and informal assessment at all levels of education such as test based or performance based assessments (Özcan & Oluk, 2007).

The three most common domains on which assessment can be made are cognitive, affective, and psychomotor skills (Bloom et al., 1956). Although affective and psychomotor skills are important in higher education as well as in any other level, cognitive skills are given more attention in admission systems. Probably, the most common classification scheme for cognitive skills is the one proposed by Bloom et al. (1956) who stated that "the use of the taxonomy can also help one gain a perspective on the emphasis given to certain behaviours by a particular set of educational plans" (p.2). Bloom uses six levels to classify cognitive performance: (1) knowledge, (2) comprehension, (3) application, (4) analysis, (5) synthesis and (6) evaluation. It enables a method to assess students for both researchers and instructors. Later, Anderson et al. (2001) proposed a revised version of Bloom's original taxonomy. As seen from Figure 1, this new classification includes more dynamic categories, which are defined by action verbs at each level: (1) Remembering, (2) Understanding, (3) Applying, (4) Analysing, (5) Evaluating, (6) Creating. The first level, 'knowledge', from the original taxonomy was replaced by 'remember' in the revised version. In the revised taxonomy, 'knowledge' is considered as the basis of the other cognitive levels. In addition, the highest two levels, synthesis and evaluation, were interchanged.

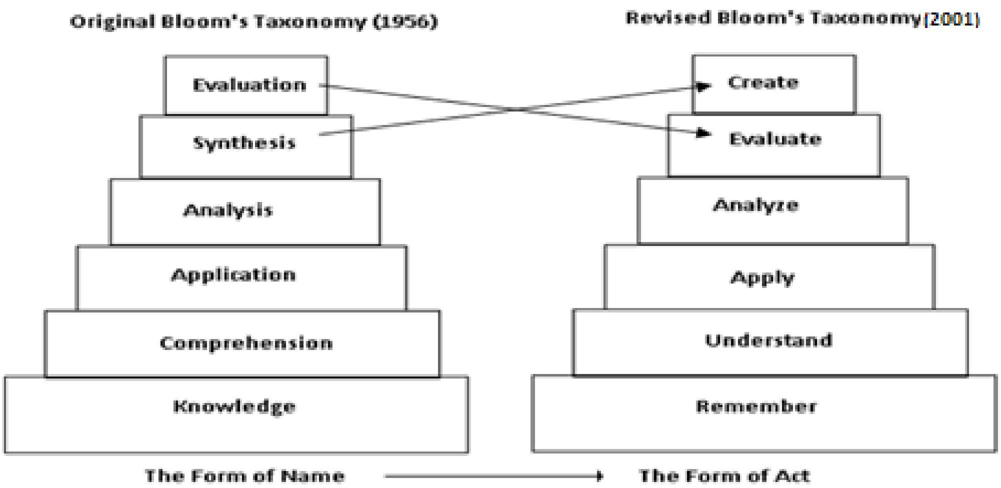

Figure 1. The comparison of Bloom's taxonomy and the revised Bloom's taxonomy

Bloom's taxonomy is utilized internationally by large-scale assessments, which have been implemented for years. For example, The International Baccalaureate (IB) Diploma Program is one of the most recognized, sufficient, and acceptable programs which is comprised of "a challenging two-year curriculum, primarily aimed at students aged 16 to 19. It leads to a qualification that is widely recognized by the world's leading universities" (Online Curriculum Centre, 2014). The examination of the IB program is found on cognitive domains of Bloom's Taxonomy (IBO, 2016).

The International General Certificate of Secondary Education (IGCSE) is an internationally recognized high school completion program. Cambridge IGCSE program is a learning process which involves verbal skills, 21st century skills, problem solving, and knowledge recollection (IGCSE, 2016). The items of this examination of the IGCSE are based on Bloom's Taxonomy.

A-level examination is an undergraduate exam that requires a sufficient score in order to be admitted into a university. The items of this examination are also prepared in accordance with Bloom's Taxonomy.

Problem

In Turkey, students are admitted into universities by a centralized and standardized admission system. The items in the HETE and the UPE mainly assess the cognitive level of students. The cognitive level of items in the tests are not only important in regards to the selection of qualified students, but also for ensuring that students who enter higher education programs are properly equipped to become critical and effective members of a country's human resource.

Thus, the present study aims to determine the level of science items in cognitive domain levels. The selection of the science subtest was chosen due to its very low means in the HETE and the UPE. In 2016, the means of science subtests was 5.03 out of 30 in physics, 9.53 out of 30 in chemistry, and 7.73 out of 30 in biology, in the HETE, respectively (MSPC, 2016).

It is known that some of the university admission examinations items can be solved by memorization rather than showing a degree of cognitive skills. Private tutoring institutions, private teachers, or preparation books provide students with clues to the correct answers of the questions. However, such clues mostly work for low-level questions. For questions assessing higher-order cognitive levels, it is much more difficult to determine the correct answer.

There are many studies in the literature detailing the analysis of the UPE questions according to cognitive levels of Bloom's taxonomy (Köğçe & Baki, 2009; Özmen, 2005; Sesli Topçu, 2007; Sönmez, Koç & Çiftçi, 2013). However, literature has an academic gap regarding the comparison between students' scores in the HETE or the UPE science subtest with their cognitive levels based on Bloom's taxonomy. The researcher initiates the study to fill this gap in the literature by assessing cognitive levels of students through their scores on science subtest besides determining the cognitive levels of UPE questions based on Bloom's taxonomy.

Purpose

One of the purposes of this study is to reveal the levels of the science items according to Bloom's taxonomy and investigate the relationship between scores and levels of items. Both the UPE items and students' science results were analysed in this study. Because Bloom's taxonomy shows effectively different levels of complexity within the questions, Bloom's Taxonomy was used by classifying science items. By this way, it is expected that a basic scheme for proficiency level to be defined for the science domain.

Research questions

This study addressed the following questions:

1) To what extent is the domain of cognitive levels of Bloom's Taxonomy covered in the science subtest in the UPE?
2) What are the difficulty levels of the items in the science subtest in the UPE?

3) What might be the cognitive levels of students who have taken the UPE based on the items they have answered correctly?

Significance

The basic function of the UPE has become a means in which to "eliminate" candidates in order to satisfy quotas set by universities, rather than measuring the competence and versatility of the students it's testing (TED, 2005). Increasing high demand to be placed in a university makes the university admission process difficult for MSPC in each year. Dependence on the UPE only ensures that all students are assessed with the same criteria for all of the programs within the universities. This study is expected to reveal information about implementation of high-stakes tests.

Moreover, students are expected to gain and improve their skills such as conducting research, observation, analysing, and evaluating the topic throughout secondary school education (Secondary schools' biology curriculum, 2013). However, the UPE consists of multiple-choice items, which is very challenging when creating a test that aims to assess students' high level thinking skills. Results of this study ca be helpful to assess use of multiple-choice items in high-stakes tests.

Prior studies have not involved students' results, but only the cognitive level of items asked in undergraduate placement examinations. There is not any study that includes students' results interpreted through the means of Bloom's Taxonomy. This study involved both the items of undergraduate placement examinations and students' scores corresponding exams.

The HETE and the UPE results are used to place students into higher education programs. However, this is not the only information that can be extracted from their results. Generally, the score of students who have taken international tests are announced with a degree that shows the ability, skill or competence of the students. For instance, international baccalaureate diploma programme's assessments are dependent on the objectives which determine students' cognitive levels. Similarly, in Turkey, each student's cognitive level can be defined based on his or her recognised cognitive levels. Similarly, higher education programs can benefit from such information when designing their curriculum by utilizing these considerations.

Students are only informed whether or not they are able to go to a university as a result of the UPE, they cannot know which cognitive level, skill or ability they have achieved. The UPE is a very extensive and demanding examination that should provide more information related to students' proficiency upon its completion.

This study may provide a description of how the university entrance exam could provide more information about students' qualifications. Students may benefit from the outcomes, if the proposed changes based on result of this research were accepted and implemented. They would have a clearer understanding of their proficiency upon learning their exam results, which in turn may direct them to concentrate on certain subjects in accordance with their abilities, even if they are not placed into a university. In addition, universities may also benefit from the outcomes; hence, they would select students according to their proficiency. The present study is expected to provide significant information as to which cognitive level, skill or ability they have.

Definition of key terms

Bloom's Taxonomy: Bloom's Taxonomy is a classification to facilitate the assessment process in education. It is a gradual order of students' skills from simple to complex. It analyses student's skills in three dimensions, which are cognitive, affective, and psychomotor domains (Bloom et al., 1956). In this study, cognitive skills of the students were studied.

Undergraduate placement examination (UPE): In Turkey, **the** University placement exam is being conducted with different names, contents and sessions every year since 1974 by Measurement, Selection, and Placement Centres (MSPC, 2014).

Difficulty parameter: Difficulty parameter is the ratio of true answers to the total number of responses (Özçelik, 1997).

CHAPTER 2: REVIEW OF RELATED LITERATURE

Introduction

The purpose of the literature review was to provide sufficient background information about university admission systems and Bloom's Taxonomy. The previous studies about the same context and their findings were explained to support the current study. In this study, the UPE items were analysed based on the domains of cognitive levels of Bloom's Taxonomy. The purpose of using Bloom's Taxonomy (cognitive domain) to analyse the items was also explained in this chapter. In addition, the features of revised Bloom's Taxonomy were given.

Assessment of students' learning outcomes

The need for validation in one's actions is a basic staple of human nature in all facets of life. When such validation derives from more coordinated work such as education, the assessments of the work done carries more importance. 'Assessment' is described as the observation of certain qualities and interpreting these observations through the use of numbers or other symbols (Turgut, 1977). In other words, assessment is performed for the purpose of determining what extent an individual obtains a certain quality or feature (Kan, 2006).

The definition of assessment varies in a particular educational context. The term of evaluation can be used instead of the term assessment in the field of education. Assessment means the judgement of students' works. Evaluation means the judgement of students' learning in the course or the action of the assessment itself. In the literature, these terms can be used interchangeably (Taras, 2005). Whereas there also exists the conflicting perception, that assessment should not be confused with evaluation. It is significant to know the difference between the concept ‘assessment’, which is the statements of definitions or symbols of any variable, and ‘evaluation’, which is interpreting symbols or descriptions obtained from the results of the assessment, achieving judgments about what is measured by comparing the criteria (Durak, 2002).

Education is the process by which an individual makes changes in his or her own behaviour through the experiences of life (Ertürk, 1972). If the behavioural changes are observable, it means that learning has occurred. One of the most important purposes of education is to train individuals to have the abilities needed to solve problems faced based on cognitive, affective, and psychomotor domains. If individuals are expected to have cognitive, affective and psychomotor abilities to solve the problems, it is not acceptable to expect the occurrence of these features to manifest haphazardly throughout course of one’s education. Thus, in order to become self-sufficient to solve the problems they will face, students should be guided by means of a specific plan. Program implementation in the framework of this plan and it may be possible to generate desired behavioural characteristics through its implementation (Alici et al., 2011). It can be concluded that education is the system to create and maintain this process.

The control of the educational system is carried out through assessment. Assessment is functioning to determine if the education system works as planned. If there is any part that becomes inoperative at any point in this system, assessment is designed to contribute to the repair in order to make the system work as planned (Alici et al., 2011).

"… all assessment begins with summative assessment (which is a judgment) and that formative assessment is in fact summative assessment plus feedback which is used by the learner" (Taras, 2005, p. 466). Burns (2008) covered that summative and formative assessment had been used with different meanings until the study of Bloom, Hastings, and Madaus. Summative tests or summative assessments are applied class-based, school-based or national, while formative assessments provide some information to students about their learning as feedback (Clarke, 2008, p. 9).

Assessment has an important role in education, which serves to illuminate the extent to which students' learning has reached a concrete stage. Students and instructors alike are informed on student's backgrounds in addition to incomplete or incorrect information as an outcome of assessment. However, in order to determine or measure what students have learned, there should be some standards set as a framework.

Black and William (1998) stated that it is essential for assessment to garner some predetermined outcomes, which are essential for the quality of the process. In order to judge and make an accurate comparison between the first and final works and thereby measure the learning outcome of students, there should be some criteria and goals by which to do so. It may be concluded that all assessment types should have such criteria and goals.

Learning outcomes

In formal education, the students are expected to gain the necessary knowledge, skills, and attitudes congruent for each grade. Arslan (2008) stated that teacher-centred approaches in education cannot be expected to be successful. Therefore, students need to be imparted with the information, attitude, and skills necessary to be accomplished within their determined level and skillset.

As a result of education, students are to gain desired qualifications, which are ultimately determined by goals that instructors or teachers set before giving a lecture. The desired goals that students can expect to gain may be at different levels and features. In the planning of assessments in education, target levels are sorted according to predetermined specifications (Varış, 1996).

Sadler (1998) supported that learning outcomes are remarkably significant for an assessment process. Black and William (1998) discussed that the achievement of students decreases when they are not provided learning outcomes when measuring skills and self-evaluation after feedback. It shows that learning outcomes are the factors that motivate students to reach set goals and furthermore, increases their achievements through these means. What and how much is expected from the students is put forth more clearly when the learning outcomes are given or targeted. In this case, according to the characteristics of the learner, improvement requires the identification and classification of targets.

Bloom et al. (1956) discussed classification of learning outcomes to facilitate teachers' measurement and evaluation of problems. It was identified that there were three domains of learning outcomes (affective domain, psychomotor domain, and cognitive domain) two of which are still used frequently by educational researchers. Affective domain included attitudes, interests, appreciation, and adjustments. Psychomotor domain aimed to assess the procedures, product, and the problem solving skills of students. Cognitive domain consists of a continuum from lower-level cognitive outcomes to higher-level thinking skills (Bloom et al., 1956). However, it is known that it is not possible to separate these domains from each other, as there is a close relationship between the two (Bakırcı & Erdemir, 2010).

Like any other theoretical model, Bloom's Taxonomy has its strengths and weaknesses. According to the researcher, the most powerful aspect is to address of a very important issue, which is thinking and is placed a structure that can be used by beginners in the profession.

Cognitive learning outcomes

According to Piaget, cognition enables us to perform the act of learning and understanding (Piaget, 1964). Cognition is a mental activity, which allow people to interpret themselves and the environment. Thus, cognitive learning may be explained by mental skills associated with the speed of mental activities.

Sorting of desired behaviours from simple to complex, from easy to difficult, from concrete to abstract in a way that is precondition is called taxonomy (Sönmez, 1997). The classification of the learning outcomes was initiated, in order to increase the efficiency of measurement of learning outcomes by Bloom et al. 1956.

Measurement of cognitive learning outcomes

Education process begin with determining the needs of individuals and the society. The desired behaviours are gained by education, which are the characteristics that society or individual want to see themselves. Therefore, the desired behaviour can vary to gain through the education process. These features needed by the members of a community constitute the objectives of education programs (Atılgan, 2006).

Students are assessed and evaluated to determine whether the desired changes appear or not. Assessment’s items should be capable of measuring desired behaviours covered in the objectives. In order to measure desired behaviours on students, each of the question on a test should depend on the fulfilment of two conditions. One of them is converting into observable behaviour by determining clear goals of teaching. Second one is selecting the type of material that suits the measurement and then carefully written to expose desired behaviour (Erman, 2008).

Some educators have attempted to form classifications of educational goals in order to express teaching objectives to be understood clearly by instructors and students. This general classification of the educational objectives as a result of Bloom and his colleagues' efforts is widely accepted. There are several reasons why this classification received well recognition. One of them is that it unifies the terminology that is used to classify and describe the learning objectives and as a result facilitates communication between educators, test developers, software developers, educational researchers and the people interested in education. Moreover, another purpose is to help educators while they are studying and comparing various training programs. The second one is to ensure taking into account all categories during teaching and assessing. The third reason is to ensure that the basic behaviours for learning are taught, before attempting to teach more difficult and complex behaviours (Atılgan, 2006).

Bloom's taxonomy

Half of the 19^{th} century classification of learning outcomes had been a source of great interest and also was subject to criticism in terms of the determination of the targets, facilitating the way to reach the targets. Despite all the criticism, it has become an indispensable part of the educational sciences. Bloom and a group of behavioural psychologist, administrators and researchers convened in 1949. The result of the decision was made in this section; a meeting was organized twice in a year for five years. They aimed to develop a classification system for cognitive, affective and psychomotor domains. The group finished their cognitive domain classification in 1956. The group's five-year study had resulted in a book named "Taxonomy of Educational Objectives. The Classification of Educational Goals, Handbook I: Cognitive Domain", which was published in 1956. The classification is called Bloom’s Taxonomy, since the leader of the group was Bloom (Krathwohl, 2002). On the other hand, affective domain classification was completed in 1964 under the title of “Taxonomy of Educational Objectives. The Classification of Educational Goals, Handbook II: The Affective Domain”. Different scientists such as Simpson (1972), Dave (1970) and Harrow (1972) developed the psychomotor domain taxonomy (Huitt, 2011).

Bloom’s Taxonomy was developed and published by a group of behavioural psychologist, administrators and researchers headed by Bloom in 1956. Bloom's taxonomy, as a theory of systematic classification in the teaching-learning process, is still widely used universally (Tutkun, 2012). Krathwohl (2002) stated that this classification was a framework to facilitate the assessment process on students’ skills.

The main purpose in construction of Bloom's Taxonomy was that the educators need to know a gradual order of student’s skills from simple to complex, and students should be able to express themselves. The levels of classification were listed consecutively. Before a top-level can be accessed, the lower level should be fully understood. (Huitt, 2011).

There are six major categories in Bloom’s taxonomy. The original Bloom’s taxonomy provides detailed and adequate definitions for each six categories in the cognitive domain. The major categories were designed in a hierarchical way from simple skills to complex skills (Krathwohl, 2002). Bloom et al. (1956) identified these categories as Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. The first three categories were accepted as low level of cognitive skills; however, the last three categories were considered as higher level of cognitive skills. In order to get to the higher steps of these skills, the previous one should be gained and digested. Any category itself is considered as a prerequisite for the next category. Bloom's Taxonomy was designed to develop a classification of educational learning outcomes to help teachers, administrators, and researchers by assessing students and evaluating learning problems (Amer, 2006).

Bloom's taxonomy, which was developed by Bloom et al., was approved at the international scale and implemented in the educational areas. However, beside this acknowledgement throughout the world, Bloom's Taxonomy has been under some criticism and some developmental studies was done to overcome these deficiencies (Arı, 2013).

Revised Bloom's taxonomy

As a result of the changes in the educational process and advances in scientific knowledge, Bloom's Taxonomy was beginning to be seen as a cornerstone in the area. The necessity for the improvement of original Bloom' Taxonomy had two main reasons. The first one was to redirect the attention on the value of the first form of the taxonomy. The second one was to improve the taxonomy based on the developments in the psychology of learning, the teaching methods and the assessment techniques (Bümen, 2006).

A revision to Bloom's Taxonomy was made and published by a group of cognitive psychologists and instructional researchers 45 years after the original book

(Anderson et al., 2001). Revised Bloom's Taxonomy was not a correction of the previous one, but a suggestion to make the classification clearer and organized for instructors in today's world.

Revised Bloom's Taxonomy was improved by creating subcategories, renaming and slightly reordering major categories. At the end of the study, although there were not fundamental changes to the classification of Bloom, it had revealed some important differences. In the study of Forehand (2005), the differences in revised taxonomy were divided into three groups. The first one was terminological changes; name of Bloom's six major cognitive categories were changed from nouns to verbs, the lowest level category knowledge was changed as “remember”, comprehension was changed as “understand”, application was changed as “apply”, and analysis was changed as “analyse”. Moreover, synthesis and evaluation taxon was renamed and replaced as “create” and “evaluate”. The second one was structural changes; while the original dimension of Bloom's taxonomy was developed in a one-dimensional frame, the revised taxonomy had two-dimensional structure as information and cognitive. The most significant change made in the revised version was to convert the one-dimensional structure of cognitive domain into two-dimensional structure (Krathwohl, 2002; Yurdabakan 2012). The third one was a purposeful change; extending the new taxonomy to a wider group in terms of the age range.

The original Bloom's Taxonomy helped further evaluation, while revised Bloom's Taxonomy focused on planning, teaching and assessment and consistency between programs. Original taxonomy helped to form the objectives of higher education, while revised taxonomy was also suitable to be used in primary and secondary school education. Original taxonomy targeted overall six major categories, while revised version focused mostly on the low cognitive levels. Original taxonomy concentrated on multiple-choice tests more (Anderson & Kratwohl, 2014). Furthermore, Amer stated (2006) Revised Bloom's Taxonomy kept aligned but improved using the recent developments made in the field of psychology and education. Since the original taxonomy was published, psychological and educational researches have continued at full speed to make students be more aware of their own thinking and learning processes and to equip students better.

Large scale assessment

The main purpose of local, national and international large-scale assessments and examinations are to inform people to what extend they have the knowledge in a particular area, to meet the requirements of the students' graduation and to develop the education programs. A wide range of examinations and assessments are performed in order to measure the skills and knowledge students had in particular areas. As a result of the large-scale assessments, students are provided the information needed in developing their skills by decision-makers in the field of education on local, national or international (British Columbia, 2013).

Large-scale assessments are convenient instruments to assess students' performance, knowledge, abilities and skills and they are implemented not only to accept students to an undergraduate program or a certifying program, but also they are efficiently used to make comparison between students, schools, and states (DePascale, 2003).

Large-scale achievement tests, in general, represent achievement tests including different level of classes or subject areas and consists of several subtests (Çakan, 2003). There are several large-scale achievement assessment tests that are done among countries to compare students' abilities, skills, literacy and creativity. TIMSS (Trends in International Mathematics and Science Study) is one of them that is conducted by International Association for the Evaluation of Educational Achievement (IEA) at four-year intervals. TIMSS is carried out to evaluate 4th and 8th grade level students in mathematics and their knowledge in science fields. In 2011, more than 60 countries participated in this study. The main purpose of TIMSS is to promote developing education and training students in mathematics and science. This project monitors trends in student achievement and determines the differences between national education systems. In the project, a great deal of information is collected about student, education programs, curricula, education systems, characteristics of students, teachers and schools. While assessing the performances of 4^{th} and 8^{th} grade students in math and science in four-year intervals, the test results also provide information about the changes occurring in student achievement levels. In this context, countries can gather detailed information on proficiency of their respective curriculum, monitor their progress as a nation, and they can compare their success in education with other countries. Besides the determination of achievement scores of students, a wide range of data is collected about schools, students and teachers through surveys. Data collected through examination and surveys provide these countries the opportunity to compare many aspects of their education system at the international level among other countries. TIMSS is an international monitoring system, which is designed in a very systematic way at all stages of the project. From the perspective of our country, TIMSS study is a very significant study in every aspect (TIMSS, 2015).

Another large-scale achievement assessment carried out in the international scale is PISA (Program for International Student Assessment) by OECD (Organization for Economic Co-operation and Development). PISA evaluates the knowledge and the skills of the students in the age group of 15. The main purpose of PISA is to get to know students better and to demonstrate students’ willingness to learn, their preferences related to learning environment, their performances in the course. PISA Project collects data from 15 age group students, who are at the last age group in the compulsory education and continues to formal education, about mathematical literacy, science literacy and reading skills, the motivation of students, their opinions about themselves, learning styles, the school environment and their families. In PISA Project, literacy concept is defined as finding, using, accepting and evaluating written resources by improving the student's knowledge and potential. In our globalizing world, policy-makers and decision-makers need to determine our rankings in the field of education on the international level, and also use those rankings for national evaluation studies (PISA, 2016).

Another large-scale assessment applied worldwide is SAT (Scholastic Aptitude Test) a test accepted in the application process of the universities in the United States and used by American citizens as well as by foreign students, also foreign students in Turkey use it to get placed in the Turkish universities. SAT is organized by The Collage Board seven times a year in the United States, and is applied six times a year in countries outside the U.S. SAT aims to measure the ability to comment literacy, composition and level of students' knowledge in the area of mathematics, geometry, which are essential in college education. SAT examination tests students’ thinking skills and knowledge, which they would need to success in academic studies in university. Usually, senior students or high school graduates take the SAT. The purpose of SAT is to measure the level of knowledge and skills acquired in secondary education. It is evaluated to assess how students solve problems, think and relate between problems. SAT is a great assessment resource for higher education institutions. This test is also one of the most sensitive indicator of the performance of students at the university.

In this context, the aim of a large-scale test must be clearly defined especially if the exam is used to make important decisions, for example, undergraduate placement examination. It is almost impossible to gather evidence for the validity of the test results of a large scale assessment, if the aim or the purposes of the exam is not well defined. For example, there are limitations in terms of exemplifying of the objectives of the national curriculum in large-scale tests in Turkey. Ministry of National Education Programs target many skills that can remain outside of the scope due to use of the multiple-choice questions in large-scale assessments.

Large scale assessment in Turkey

University admission system is a large scale assessment. “The term large scale assessment refers to any provincial, national or international assessment, examination or test the Ministry directs boards of education to administer” (British Columbia, 2013).

During the Republican period, universities admitted the high school graduates without examination until 1960. Faculties that exceeded their quota usually made their choices by one of the following ways:

•A first-come first-serve method where they accepted applicants until their quota was met.

•According to the quality of education of the high school graduates in science or literature.

•High school degree of high school graduate students and accept students by sorting their degrees (MSPC, 2008).

As the amount of applications to higher education increased, student selection methods outlined above could not meet the needs; faculties, according to their purpose, began to organize entrance exams. In the latter case, students had to scour from city to city to participate in these exams. Students had to face circumstances where they had two exams on the same day in two different cities. This situation had considerably led to complaints from both candidates and their parents.

In the 1960s, firstly, some universities began to give entrance exams by themselves; later some universities attempted to act together in order to give a common exam. The increase in the number of candidates for these exams required the utilization of informatics methods and tools such as reporting results, applying, scoring, selection and placement. In 1974, the university entrance exam was decided to be done by a single centre, hence the Inter-University Council and MSPC (Measurement, Selection and Placement Centre) was established for this purpose.

Until 1981, the selection and placement of students to the university was conducted by the MSPC. In 1981, MSPC turned into a subsidiary of the Board of Higher Education. In 1974 and 1975, Student Selection and Placement Examination (SSPE) was given in two sessions during morning and afternoon on the same day.

In 1976-1980, it was given in one session; till 1981, it was applied in two-session state exams. In the first step, the two-stage examination (SSE) was implemented in April. According to the results of the first session, the potential candidates were identified and given the right to take the second exam, which was the SPE given in

June. Undergraduate placement examination was carried out in two sessions between 1981 and 1999 (MSPC, 2008).

With the changes made to the undergraduate placement examination to be administered in a single-session in 1999, the scope of subjects in the examination was limited; and the questions were simplified (Özmen, 2005). The SSE continued to be given in one session with the amendment in 2006 until 2009. Before 2006 SSE test questions had been prepared usually in primary and secondary acquisition of basic knowledge and skills based on students' reading comprehension, reasoning, relationship building, and so their abilities were measured according to this principle.

In 2006, this system was changed. Exam questions included learning curriculum from the second and third grade of secondary school students (MSPC, 2008). Moreover, some of the questions were kept similar to the ones in the previous year, while some questions were prepared taking into consideration the entire high school curriculum (MSPC, 2014).

It takes a long process to prepare questions. Firstly, it is determined in which extent the questions will cover the program of the high school curriculum. At this stage, it is important to determine whether or not the questions had been asked in the past or if they are similar to the questions in the recent books. Then a draft is created. After it is agreed on the cognitive level of the questions, the committee determines the final versions of the questions. Moreover, it is important to obey the test rules, for example whether or not the information on the questions are given in the books taught in the schools or the answers of the items are clear enough (Erman, 2008).

Studies on Bloom's taxonomy in undergraduate placement examination

Multiple-choice exam is an instrument for education, it is not a purpose. It is possible to identify how well teaching and learning activities succeeded and based on these necessary arrangements can be made in the curriculum according to the scores of students in large scale assessments. According to Berberoğlu and İş Güzel (2013), some studies can be done on comparing the achievement levels of students from year to year or the failure of the education system based on the low achievement levels of student. However, it is not reliable to compare results different years.

Many studies were carried out in order to analyse the questions asked in undergraduate placement examination according to Bloom's Taxonomy. For example, Sesli Topçu (2007) stated that considering the cognitive levels of biology questions in UPE, application level (36%) was asked the most; following this analysis level (35%) and synthesis level (23%) were observed in classification of the levels of the questions.

It was determined that math questions asked in the UPE were mostly application level (Köğçe & Baki, 2009). Özmen (2005) stated that the UPE chemistry questions were mostly at the application level, following this understanding and analysis questions were asked mostly. There have been some studies in social science that indicated similar results. Distribution of the UPE geography questions based on the cognitive domain of Bloom's Taxonomy was concentrated mainly at the level of understanding (Sönmez, Koç & Çiftçi, 2013).

According to the studies mentioned above, students were asked questions mostly at the middle cognitive levels of Bloom's taxonomy. However, according to experts UPE items should be prepared in order to prompt students' higher order thinking levels (Karamustafaoğlu, Sevim, Karamustafaoğlu & Çepni, 2003; Tezbaşaran, 1994).

Bakırcı and Erdemir (2010) found out that first year science pre-service teachers were in low cognitive levels of Bloom's Taxonomy, while last year pre-service teachers had higher thinking skills. This shows students were improving their higher thinking skills throughout their education in university. The reason is that students had to attend private tutoring centres to succeed in the UPE and to get used to the multiple-choice exams. However, private tutoring centres focus on memorization of knowledge which cause students to have low cognitive level skills (Bakırcı & Erdemir, 2010).

Özmen and Karamustafaoğlu (2006) studied high school physics and chemistry exam questions, besides UPE exam, according to the analysis of cognitive levels. The result of the study showed that students generally respond correctly to the questions in knowledge, comprehension and application levels in high school physics and chemistry exams.

Akpinar and Ergin (2006) studied written science examination questions at certain high schools. The results of the research showed that most of the questions asked in the written exams in the field of science were in knowledge and comprehension level in schools. Moreover, it was concluded that teachers assessed whether their students memorize the information or not, they did not compel their students to analyse, synthesize or evaluate the knowledge.

Kadayıfçı (2007) investigated 210 examination papers of 42 chemistry high school teachers in another study based on Bloom's Taxonomy. In the result of the questionnaires that applied to the teachers, teachers were saying that they included the questions from all six cognitive levels of Bloom’s Taxonomy. However, it was concluded that the questions of the examinations were in the first three cognitive levels of Bloom and the situation seemed to be related with the level of UPE chemistry questions.

Azar (2005) studied physics high school exam questions and UPE physics questions on Bloom's Taxonomy. In this study, 76 questions in physics from 2000-2003 UPE and 556 physics questions of 12 physics high school teachers at the same year were collected and analysed according to Bloom’s Taxonomy. It was found that UPE consisted of questions measuring application, analysis, synthesis and evaluation skills of students, while physics exam questions were at the knowledge, comprehension and application level of Bloom’s Taxonomy.

Köğçe (2005) examined UPE items from 1995 to 2004 and math questions asked in written examination at high schools from 2003 to 2005 to compare them according to revised Bloom Taxonomy. Total of 2300 questions were analysed in accordance with the criteria prepared taking into consideration the cognitive level of revised Bloom’s Taxonomy. As a result of this research, Vocational High School and Technical High School teachers tended to ask comprehension cognitive level of questions, while Anatolian and Science High School teacher were found to ask questions mostly on the application and analysis levels. It was found that questions asked in UPE were mostly in understanding and applying cognitive level of revised Bloom’s Taxonomy. At the end of the study, it was concluded that including higher thinking cognitive level questions in school examinations will help students to think critically, analyse and evaluate the questions more.

Karamustafaoğlu et al. (2003) investigated the analysis of 403 high school chemistry exam questions collected by high school teachers based on Bloom's Taxonomy. It was concluded that 96% of the questions asked in the exams were measuring the lower-order thinking skills.

Therefore, the studies stated above shows that the UPE, textbooks and school exams mostly address lower cognitive levels according to Bloom’s Taxonomy. Thus, the items in these examinations assess only low cognitive levels of the students. In this case students are not assessed by taking into consideration their analytical, critical thinking, creativity, innovation, problem solving, informatics, teamwork, entrepreneurship and responsibility, which are the expected skills and must be gained throughout the science education in high school.

CHAPTER 3: METHOD

Introduction

In this chapter, firstly, research questions, research design and the context are presented in order to give the overall summary of the study. Secondly, instrumentation, sampling, method of data collection and method of data analysis are covered.

The following research questions are investigated in this study;

1) To what extent is the domain of cognitive levels of Bloom's Taxonomy covered in the science subtest in the UPE?
2) What are the difficulty levels of the items in the science subtest in the UPE?
3) What might be the cognitive levels of students who have taken the UPE based on the items they have answered correctly?

Research design

This study was envisioned as a quantitative exploratory research. In exploratory research, researchers examine an issue in which they do not have enough information about the topic before the investigation. This type of research provides relevant background information to the researchers and aims to gather superficial information. This kind of research is preferred in situations such as; 1) No or few study has been done about the desired group, process, activity or situation 2) The topic is resiliently unexamined by depicting 3)The topic has undergone enough changes to override the results of the study 4) It is used in order to resolve the uncertainties in the research problem and to limit the topic 5) If it is needed, the method of the study is changed according to the type of information and resources encountered (Neuman, 2000; Stebbins, 2001; Struwig & Stead, 2001).

Context

Thousands of individuals take the UPE each year. In 2007, the year of the exam this study is based on, there was a two-stage examination. A total number of 1.615.534 students took UPE in 2007. Only 392.657 of these students were placed in higher education programs (MSPC, 2007). Such a large group covers a broad range of students in terms of their abilities, cognitive skills and school types.

According to the document published by MSPC (2014), the last regulation on the UPE was made in 2006 in terms of the content. UPE items have been implemented with the same content and in one session since 2006. Even though the number of session was one; the exam had two different booklets with two different contents. First booklet was prepared to test analytical thinking skills of students, however second booklet included all high school curricula. In this study, the researcher used the first booklet of the science subtest of the UPE items from 2007.

In 2007 UPE consisted of two different booklets. The first booklet had the Turkish language subtest (30 items), the mathematics subtest (30 items), the science subtest (30 items) and the social subtest (30 items). The second booklet had the same number of each subject. All of the candidates were supposed to answer the items in the first booklet. In the second booklet, candidates answered the subtests, which were weighed the test score according to the higher education program they preferred. Students have 195 minutes to answer these tests (MSPC, 2007).

There was a total of 30 items in science subtest among which thirteen were in physics, nine were in chemistry and eight were in biology subjects in the first booklet. Items consisted of multiple-choice question. MSPC has right to change the way they ask questions from year to year (MSPC, 2007). The purpose of the science subtest was to probe the basic concepts and principles of thinking skills (MSPC, 2010).

In the first stage of the study, the science subtest of UPE items are examined and categorized based on Bloom's Taxonomy. Bloom et al. investigated the domain of cognitive levels under six different categories. These are Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation (Bloom et al., 1956). This systematic classification is known as Bloom's Taxonomy. Bloom's Taxonomy is widely used in the classification of educational objectives by teachers, instructors and educational program developers. The main feature of Bloom's Taxonomy is to help the teacher to find an answer to the question; "What kind of cognitive changes are observed in students throughout their education process?" (Küçükahmet, 2005). The cognitive domain of Bloom's Taxonomy is a level which is related to knowledge that students already have. Teachers or instructors have to ask questions to make understand these cognitive behaviours (Bacanlı, 2005).

Instrumentation

The purpose of UPE is to place students in higher education programs according to their abilities, skills and interests. Thus, in this study the items asked in UPE are classified based on Bloom's Taxonomy in order to observe students' cognitive levels when they are placed in higher education programs.

The requirements used in this study are given below for the classification of the six categories in Bloom's Taxonomy (Bloom, 1956; Krathwohl, 2002; Bakırcı & Erdemir, 2010).

Table 1
The cognitive domains of Bloom's taxonomy

Cognitive Domain	**The keywords characterised cognitive learning outcomes**
Knowledge: Students should be able to remember and recognize the information without understanding deeply.	***E.g.*** arrange, define, indicate, identify, match, memorize, list, name, label, point, recall, repeat, review, state, outline, order, select, show, write.
Comprehension: Students should be able to understand the topic without associating it with any information that is taught, and interpret the information. Students are expected to make a prediction about further questions.	***E.g.*** classify, contrast, compare, describe, differentiate, distinguish, discuss, estimate, express, extend, generalize, give, interpret, observe, predict, rewrite, relate, summarize.
Application: Students should be able to use their previous knowledge, choose the appropriate one and apply the relevant information based on the question. At the end of the question, students are expected to come up with the results.	***E.g.*** apply, assign, calculate, change, capture, classify, compute, construct, explore, draw, figure, graph, illustrate, manipulate, modify, project, solve, show, use.
Analysis: Students should be able to develop the ability of breaking the material into small pieces to understand the whole structure, and understand the cause-effect relation in the material. Students are expected to express the meaning of each piece and to analyse the relations between these pieces.	***E.g.*** analyse, break down, confirm, classify, correlate, detect, diagnose, graph, factor, illustrate, discover, determine, discriminate, examine, explore, handle, investigate, modify, predict, solve, transcribe, translate.

Table 1 (con't)
The cognitive domains of Bloom's taxonomy

Synthesis: Students should be able to create a new idea from their previous observations and experiences by combining different information.	***E.g.*** assemble, abstract, arrange, categorize, code, combine, compose, create, develop, devise, enhance, explain, formulate, improve, model, organize, outline, produce.
Evaluation: Students should be able to interpret and draw conclusion about the information by using results based on scientific truth and validity.	***E.g.*** appraise, assess, conclude, criticize, defend, estimate, evaluate, interpret, judge, justify, measure, test, support, verify.

Sampling

The science items of year 2007 are selected due to availability of its data set. Researcher accessed students' 2007 UPE scores from MSPC. Random sampling is used to form a sample for the study. According to Urbach (1989), random sampling should represent the miniature of the population or the characteristics of the population in which the study is interested. It other word the sample, which is going to be used for the study, should have the same ratio for the investigated phenomenon as the total population. If it is impossible to study the whole population or the population number is very big for the study, the group of people who represent the total population can be studied. In this sense, random sampling is a nice tool to explain an existing phenomenon (Shewhart, 1931).

Data set included randomly selected 3382 students, who took UPE while they were still at the last grade of high school.

Method of data collection

Data set of the UPE science subtest items was provided by MSPC's database. The items are classified by three pre-service science teachers according to Bloom's Taxonomy. In order to classify 30 science items, raters looked for active verbs of Bloom's Taxonomy in the roots of the items (see Table 1).

The data including student's UPE scores were also provided by MSPC's database. Following methods are followed in the calculation of standard scores of each candidate;

- The number of true and false answers given by each candidate is determined.
- The mean and the standard deviation of the test are found by using the raw scores of the candidates who took the exam the year before.
- The mean and the standard deviation of the test were found by using the raw scores of the candidates who were studying in the last year.
- Eventually, standard scores are calculated for all candidates by using the mean and the standard deviation (MSPC, 2006)

In this study, items that are answered were taken into consideration in order to find students' scores. However, unanswered items are ignored in order to avoid taking into account the students' scores, who did not take the exam.

Method of data analysis

Firstly, in order to explore the domain of cognitive levels of Bloom's Taxonomy covered in science items in the 2007 UPE; the researcher identified the cognitive levels of 30 science items asked in the 2007 undergraduate placement examination according to Bloom's Taxonomy which consists of Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation levels. In order to increase the reliability of this classification, three pre-service science teachers were asked to identify the science items as raters.

The three raters were qualified enough to determine the level of questions with respect to Bloom's Taxonomy. The three raters classified items without knowing each other's decisions at different times and places. After, the researcher collected all the classified items, a deliberation section was organized and all the raters together with the researcher had a chance to discuss their decisions. In the very beginning of the discussion, some of the decisions about the cognitive levels of items were different. Although there were disparities between some of the decisions on the levels, the levels chosen by the raters were still close enough to each other in terms of taxonomic levels of Bloom. Knowledge, Comprehension and Application level of questions are sort of low levels of Bloom's Taxonomy. However, Analysis, Synthesis and Evaluation levels of questions are considered as higher levels of Bloom's Taxonomy. For instance, if the level of one item was decided as application level by Rater-1 then Rater-2 decided as comprehension level of Bloom's Taxonomy for the same item. The differences are not very significant for those two levels, because the items are considered in the low cognitive levels of Bloom's Taxonomy.

The inter-rater reliability is calculated in Statistical Package for the Social Sciences (SPSS). It uses statistical methods to measure the degree of agreement between raters (Gwet, 2001). Inter-rater reliability provides a way to find a numerical value about a qualitative assessment that way it can be converted to a quantitative result.

After identifying the level of the questions, students' scores were statistically analysed for each question in SPSS. Difficulty parameter of each science item was analysed in SPSS and a related table was created and represented in the results.

Difficulty was considered as a parameter which is defined as the ratio of true responses to the total number of response (Özçelik, 1997). It means if the total number of students are 32 who answer an item, but the number of responses are 16 in this test, then the difficulty parameter is 50%. The distribution of the cognitive levels of the items does not indicate the difficulty of the items. But the correlation between this parameter and the true answers to each level is investigated.

Cognitive levels of the students were analysed and interpreted in accordance to the number of the true answers they had given based on Bloom's Taxonomy. Students were divided into six groups according to the number of true answers they had given;

1) Students who answered 0-5 items correctly
2) Students who answered 6-10 items correctly
3) Students who answered 11-15 items correctly
4) Students who answered 16-20 items correctly
5) Students who answered 21-25 items correctly
6) Students who answered 26-30 items correctly

According to these groups, students are expected to answer correctly to certain levels of Bloom's Taxonomy. For example; for group-1, students answered 0-5 items, are expected to correctly answer mostly to knowledge cognitive level items. For group-6, students who answered 26-30 items correctly, were expected to correctly answer both low level and high cognitive level of items. At the end of the analysis the cognitive levels of the science items and related scores of these groups are discussed based on Bloom's Taxonomy.

This study helped to describe students' cognitive levels based on classification of the UPE items according to Bloom's Taxonomy. It is unravelled whether or not the students who got placed into a higher education program have higher-order cognitive skills.

CHAPTER 4: RESULTS

Introduction

In this chapter, the researcher analysed, found and represented the domain of the cognitive levels of science items covered in the 2007 UPE with regard to Bloom's Taxonomy. In addition to that, distribution of difficulty parameter of science subtest was presented. Moreover, the inter-rater reliability was found. Ultimately, according to the students' scores in the 2007 UPE the cognitive levels of students were showed.

The cognitive levels of the items in the science subtest of the UPE

The first goal of this research was to determine "To what extent the domain of cognitive levels of Bloom's Taxonomy was covered in science items in the UPE.". In this context, Table 2 shows the cognitive level of each item as determined by four raters including the researcher's responds. Table 2 shows raters' first responses in evaluating the taxonomic level of 30 science items. The last column of the table shows the unanimous decision made by all raters. Agreement column shows whether there was an agreement or not among raters before the deliberation. The final decision after discussion of the raters is presented in the last column. As seen in the Table 2, there were 20 items which the raters concurred on assigning them to the same cognitive level and 10 items on which they disagreed. This corresponds to 66.67% agreement and 33.33% disagreement between raters about the cognitive levels of items with respect to Bloom's Taxonomy. Even if there was a disagreement about the cognitive level of an item, the cognitive levels of those items were generally low. To give a brief example, item four and six were decided to be on application level and comprehension level by different raters, however, these two cognitive levels are still in the lower level of Bloom's Taxonomy. Item seven, eight, 10, 16 and 30 were considered to be on application and analysis levels. Analysis level is sort of higher thinking level, while application is low order thinking level according to the Bloom's Taxonomy. Nevertheless, these two cognitive levels are still close in this cognitive classification. Item 17, 21 and 26 were classified in knowledge and comprehension levels, which are low levels of the Bloom's Taxonomy.

Table 2
Analysed items according to Bloom's taxonomy

Item no	Rater-1	Rater-2	Rater-3	Rater-4	Agreement	Agreed Level
1	Comp	Comp	Comp	Comp	Agreed	Comp
2	Comp	Comp	Comp	Comp	Agreed	Comp
3	App	App	App	App	Agreed	App
4	App	App	App	Comp	Disagreed	App
5	App	App	App	App	Agreed	App
6	App	App	App	Comp	Disagreed	App
7	App	Anal	App	App	Disagreed	App
8	App	Anal	Anal	App	Disagreed	Anal
9	App	App	App	App	Agreed	App
10	App	Anal	Anal	App	Disagreed	Anal
11	Anal	Anal	Anal	Anal	Agreed	Anal
12	App	App	App	App	Agreed	App
13	App	App	App	App	Agreed	App
14	Comp	Comp	Comp	Comp	Agreed	Comp
15	Comp	Comp	Comp	Comp	Agreed	Comp
16	App	App	Anal	App	Disagreed	App
17	Know	Know	Comp	Comp	Disagreed	Comp
18	Know	Know	Know	Know	Agreed	Know
19	Comp	Comp	Comp	Comp	Agreed	Comp
20	Anal	Anal	Anal	Anal	Agreed	Anal
21	Know	Know	Comp	Comp	Disagreed	Comp
22	App	App	App	App	Agreed	App
23	Comp	Comp	Comp	Comp	Agreed	Comp
24	Know	Know	Know	Know	Agreed	Know
25	Know	Know	Know	Know	Agreed	Know
26	Comp	Know	Comp	Comp	Disagreed	Comp
27	Anal	Anal	Anal	Anal	Agreed	Anal
28	Comp	Comp	Comp	Comp	Agreed	Comp
29	Know	Know	Know	Know	Agreed	Know
30	App	App	Anal	App	Disagreed	App

Note: Know: Knowledge Comp: Comprehension App: Application Anal: Analysis

Table 3 shows the analysis of inter-rater reliability. The inter-rater reliability is found to be around 0.9 for this study. This means statistically speaking there is no significant difference between the decisions made by raters, since the value of inter-rater reliability is very close to one. According to the result, the reliability of the raters was high enough to proceed the research.

Table 3
Inter-rater reliability

	Intra-class Correlation	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	0.900	0.834	0.946	39.545	29	87	< .01
Average Measures	0.973	0.953	0.986	39.545	29	87	< .01

Table 4 shows that science subtest consisted of 13.33 % (n=4) knowledge level, 33.33% (n=10) comprehension level, 36.67 % (n=11) application level, and 16.67% (n=5) analysis level of items. The first three levels are low levels of the Bloom's Taxonomy. Thus, the 83.33% of the total items are comprised of low level items. It can be concluded that science subtest generally consisted of low cognitive level items based on Bloom's Taxonomy. Only 5 out of 30 science items were observed to measure the higher-order thinking skills of the students at analysis level.

Table 4
Numbers and percentage of items according to Bloom's taxonomy

Level of Bloom's Taxonomy	Item number	Number of item	Percentage
1 Knowledge	18, 24, 25, 29	4	13.33%
2 Comprehension	1, 2, 14, 15, 17, 19, 21, 23, 26, 28	10	33.33%
3 Application	3, 4, 5, 6, 7, 9, 12, 13, 16, 22, 30	11	36.67%
4 Analysis	8, 10, 11, 20, 27	5	16.67%
5 Synthesis	-	0	0%
6 Evaluation	-	0	0%

The first 13 items between numbers 1-13 are physics items, the next nine items between numbers 14-22 are chemistry items and the rest of eight items between numbers 23-30 are biology items in this science subtest.

Table 5 displays the cognitive domain of Bloom's Taxonomy of the science items by comparing them among science subtests. When different science subtests are analysed, it is clearly seen that the distribution of the cognitive levels over items varies among the subjects.

Table 5 indicates that there is no knowledge level item in physics subtest. In the physics subtest, the most frequent type of cognitive level that appears in items is application level (n=8). The next level that appears the most is analysis level (n=3) and the one that appears the least is comprehension (n=2). In chemistry subtest, there are mostly comprehension level (n=5) items, secondly application level (n=2) items, and one item for each knowledge (n=1) and analysis levels (n=1). In biology subtest, the most asked level of item is knowledge (n=3) and comprehension levels (n=3). Application and analysis level items are asked only once for each (n=1).

It is particularly notable that the physics subtest has an application cognitive level of items (61.54%) more than that of the other subjects. The biology subtest has the lowest (knowledge) cognitive level of items (37.50%) more often than other subjects do with regard to Bloom's Taxonomy.

Table 5
Levels of Bloom's taxonomy of different subtests

	Knowledge	Comprehension	Application	Analysis
Physics	0 (0%)	2 (15.38%)	8 (61.54%)	3 (23.08%)
Chemistry	1 (11.11%)	5 (55.56%)	2 (22.22%)	1 (11.11%)
Biology	3 (37.50%)	3 (37.50%)	1 (12.50%)	1 (12.50%)

The difficulty levels of the items in the science subtest of the UPE

Table 6 represents the percentages of true, false and unanswered items. Difficulty parameter is defined as the ratio of true answers to the total number of responses for this study. Thus, the unanswered items cause decrement in difficulty parameter of the item, which causes the item to be perceived more difficult than it may be in reality. In Table 6, in calculating percentages the researcher takes into account all the items: true, false, unanswered. P column shows the difficulty parameters of the items, when the unanswered items are not counted towards the calculation to get a realistic result. Thus, difficulty parameter under P column shows the percentage of true responses out of the sum of true and false responses. Table 6 provides all the information about percentages of the items.

Table 6
The difficulty parameters of the items

Items	True	False	Unanswered	P	Items	True	False	Unanswered	P
1	9.9	22.6	67.5	30.5	16	14.0	15.8	70.2	47.1
2	15.6	20.8	63.5	42.8	17	19.2	17.6	63.2	52.2
3	13.8	24.0	62.2	36.6	18	30.7	12.0	57.3	71.8
4	11.4	22.8	65.8	33.4	19	14.3	16.6	69.1	46.1
5	18.5	23.6	57.9	43.9	20	8.3	15.4	76.2	35.1
6	22.0	15.9	62.0	58.0	21	14.9	14.3	70.8	51.1
7	15.5	23.4	61.1	39.8	22	16.8	14.0	69.2	54.5
8	19.8	23.6	56.5	45.6	23	19.6	28.6	51.8	40.7
9	22.5	21.5	55.0	51.2	24	14.3	18.8	67.0	43.2
10	22.5	15.3	62.2	59.5	25	13.4	30.0	56.7	30.8
11	28.4	19.9	51.7	58.8	26	17.7	20.7	61.6	46.1
12	22.0	18.7	59.3	54.0	27	9.7	20.8	69.5	31.7
13	11.2	20.0	68.9	35.9	28	24.4	21.0	54.6	53.7
14	16.7	18.8	64.5	47.1	29	33.3	16.4	50.3	67.0
15	14.7	16.1	69.2	47.8	30	19.4	14.4	66.2	57.5

In order to compare the difficulty parameters of items belonging to different cognitive levels, Table 7 is created. Each item is categorized as knowledge, comprehension, application and analysis according to their levels in terms of Bloom's Taxonomy. The difficulty parameters of the items are found by calculating the value of true items as one and false items as zero.

Clearly, the difficulty parameter of the item that is answered correctly the most (item number 18) is 71.8 %. This means that almost seven out of 10 students gave the right answer for this item so the difficulty parameter of this item is low. However, the smallest difficulty parameter is 30.5% for item 20, which shows that only three out of 10 students gave the true response for this item.

Table 7
The difficulty parameters of the items based on Bloom's taxonomy

Item No	Level of Bloom Taxonomy	P	Item No	Level of Bloom Taxonomy	P
25	Knowledge	30.8	13	Application	35.9
24	Knowledge	43.2	3	Application	36.6
29	Knowledge	67.0	7	Application	39.8
18	Knowledge	71.8	5	Application	43.9
1	Comprehension	30.5	16	Application	47.1
23	Comprehension	40.7	9	Application	51.2
2	Comprehension	42.8	12	Application	54.0
19	Comprehension	46.1	22	Application	54.5
26	Comprehension	46.1	30	Application	57.5
14	Comprehension	47.1	6	Application	58.0
15	Comprehension	47.8	27	Analysis	31.7
21	Comprehension	51.1	20	Analysis	35.1
17	Comprehension	52.2	8	Analysis	45.6
28	Comprehension	53.7	11	Analysis	58.8
4	Application	33.4	10	Analysis	59.5

Table 7 shows that difficulty parameter of the knowledge level items varies in a broad range compared to the other levels. To give an example, the difficulty parameter of item 25 is 30.8 whereas for item 18 it is 71.8. For the comprehension level items, it ranges from 30.5 to 53.7. Table 7 indicates that for the application level items the difficulty parameters are between 33.4 and 58.0. And for the analysis level items it ranges from 31.7 to 59.5. Therefore, it is hard for students to determine their cognitive levels according to this science subtest.

According to the Table 7, the item that was answered correctly the most (item 18), which is in knowledge level of Bloom’s Taxonomy, is one of the easiest item. The most difficult item (item 20), which is rarely answered correctly, is in comprehension level. On the other hand, students answered some of the items in analysis level mostly correct, for instance item number 10.

We observe that the low difficulty parameter of an item does not mean that the cognitive level of this item is in low order thinking level as well. In other words, the items do not have to be in higher order cognitive level in order to be considered difficult. Even if the cognitive level of an item is low, the difficulty parameter for this item can be high, because difficulty parameter is measured by the percentage of true responds over the total responds given by students.

As seen in Figure 2, the most of the items were in lower levels of Bloom’s Taxonomy (Knowledge, Comprehension, and Application) in this science subtest. The difficulty distribution of items shows that the items are mostly difficult. Even though the items were in low cognitive levels based on Bloom’s Taxonomy, we notice that items were hardly answered correctly by students.

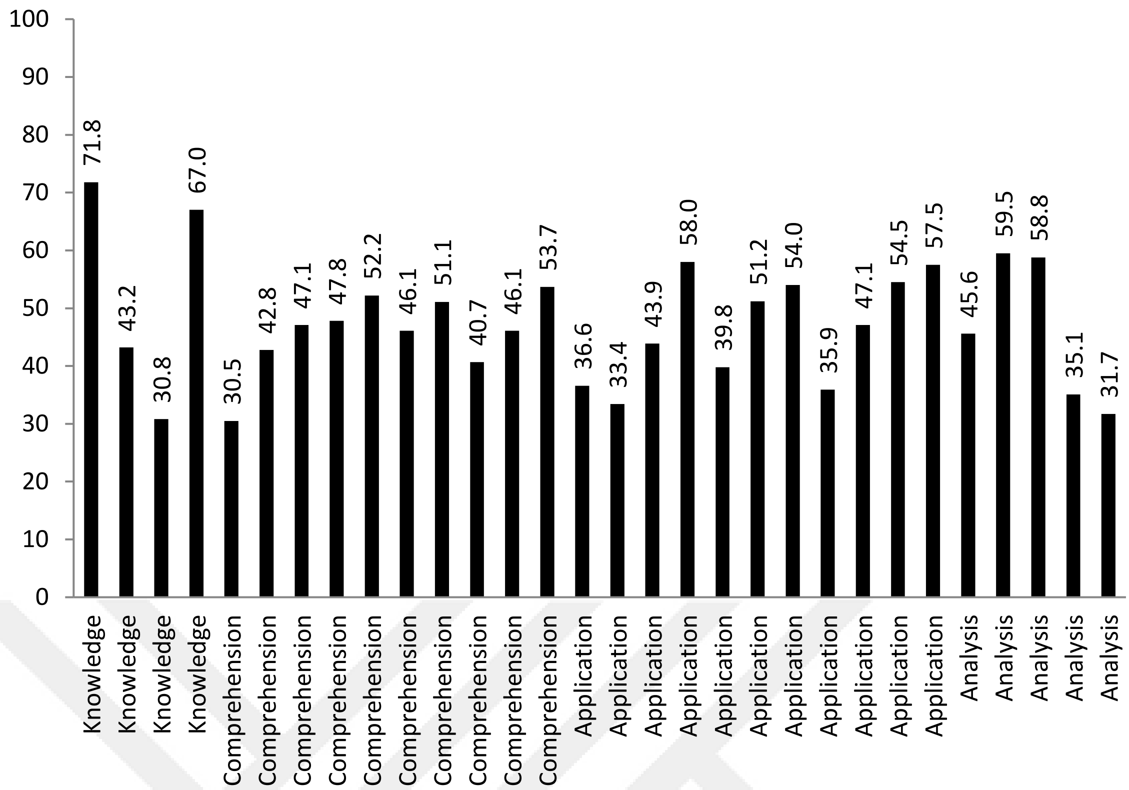

Figure 2. Difficulty parameters of items based on cognitive level of items

Table 8 indicates the difficulty parameters of each item which are grouped according to their subject. Out of 100 students 35.5 responded correctly to the most difficult physics item, which is number-1 in comprehension level. The easiest item is number-10 in analysis level, which is responded correctly by 59.5 out of 100 students. In chemistry, the most difficult item is number-20 in analysis level, which is responded correctly by 35.1 out of 100 students, while the easiest item number-22 in application level, was responded by 54.5 out of 100 students. In biology subtest, the most difficult item is number-25 in knowledge level, which is correctly responded by 30.8 out of 100 students. The easiest item number-29 is again in knowledge level and is correctly responded by 67.0 out of 100 students.

As seen from Table 8, the difficulty levels of physics items (SD= 9.7) do not show much variation, as well as chemistry items (SD=9.7). The difficulty levels of biology items vary in a broader range than those of physics and chemistry.

Table 8
The difficulty parameters of items based on subject area

Number of item	Physics	Number of item	Chemistry	Number of item	Biology
1	30.5	14	47.1	23	40.7
2	42.8	15	47.8	24	43.2
3	36.6	16	47.1	25	30.8
4	33.4	17	52.2	26	46.1
5	43.9	18	71.8	27	31.7
6	58.0	19	46.1	28	53.7
7	39.8	20	35.1	29	67.0
8	45.6	21	51.1	30	57.5
9	51.2	22	54.5		
10	59.5				
11	58.8				
12	54.0				
13	35.9				
Average	45.4		50.3		46.3
SD	9.7		9.7		11.7

Figure 3 represents difficulty parameters of items based on subject area. It can be deduced that the distribution of the difficulty levels in each subject differs from subject to subject but is relatively homogeneous in each subject.

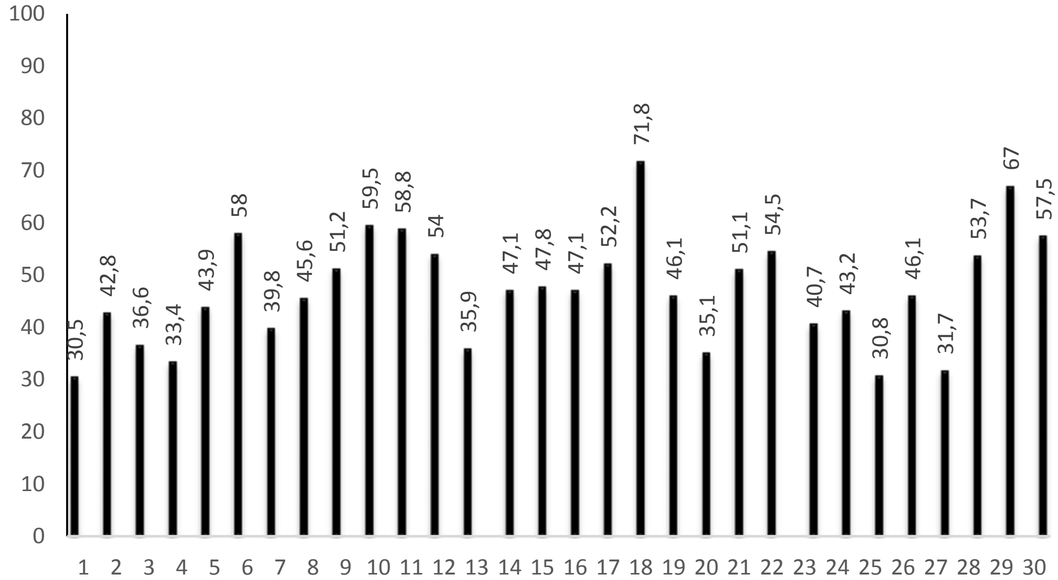

Figure 3. Difficulty parameters of items based on subject area

The descriptive statistics of difficulty parameters of the true items are calculated and shown on Table 9. The mean of the true items is found to be 47.12 (SD=6.05). While chemistry subtest has the easiest item (P=71.8) in the test, physics has the most difficult one (P=30.5).

Table 9
The descriptive statistics of difficulty parameter of true items

True items	Science	Physics	Chemistry	Biology
Mean	47.12	45.38	50.31	46.34
Median	46.60	43.35	47.80	44.65
Mode	46.10	-	47.10	-
Std. Deviation	10.53	9.68	9.71	11.74
Minimum	30.50	30.50	35.10	30.80
Maximum	71.80	59.50	71.80	67.00

Figure 4 shows the difficulty distribution of the items. Mode having a lower value than mean and median shows that difficulty distribution is right-skewed, which is indicating that the items were relatively difficult.

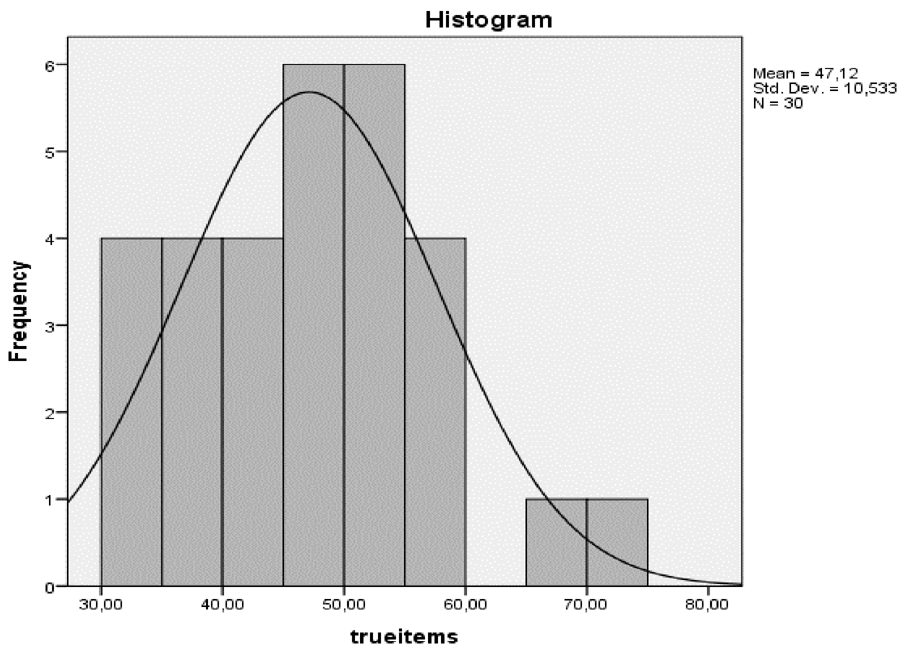

Figure 4. Difficulty distribution of true items

Table 10 shows what percentage of items in each cognitive level is correctly responded by students. For instance, the total number of students who answered knowledge level items is 907. This means, 2.10 (SD=1.27) out of 4 items are correctly answered in knowledge level, which corresponds to 57.5 % of the total number of knowledge level items. Comprehension level items are correctly answered by 553 students. This is 4.60 (SD=3.21) out of 10 items, which is equal to 46.0 percent of total comprehension level items. Moreover, 517 students correctly answered to application level items. This corresponds to 4.97 (SD=3.41) items out of 11 application items, which is equal to 45.18 percent of total application level items. For the analysis level items, 2.25(SD=1.59) out of 5 items are correctly answered by 641 students which corresponds to 45.0 percent of total number of analysis level items.

When the results of the answered items are compared, it is clear that the knowledge level items are the ones that got mostly true responses (52.5%), the comprehension level items came second (46.0%), the application level items are correctly answered third (45.18%), the analysis level items are the fewest correctly answered items (45.0%). This comparison shows that students tend to answer correctly to low cognitive level items than higher thinking level items based on Bloom's Taxonomy.

Table 10
Total number, mean and standard deviation with respect to Bloom's Taxonomy

	N	# of items	Mean	SD	%
Knowledge	907	4	2.10	1.27	52.50
Comprehension	553	10	4.60	3.21	46.00
Application	517	11	4.97	3.41	45.18
Analysis	641	5	2.25	1.59	45.0

Note: N: Total number of student, SD: standard deviation

The cognitive levels of the students based on science subtest of the UPE

Table 11 represents the overall success percentile of students. As seen from Table 11, out of 3382 students 2206 (65.23 %) of them got the true responses to only 0 to 5 items. Only 100 (2.95%) students correctly answered more than 26 items.

Table 11
The success percentile of the students

# of true responses	Number of students (n)	Percent of Students
Between 0-5	2206	65.23 %
Between 6-10	574	16.97 %
Between 11-15	233	6.89 %
Between 16-20	132	3.90 %
Between 21-25	137	4.05 %
Between 26-30	100	2.95 %
Total	3382	100 %

Table 12 shows the success rate in each cognitive level of the students who correctly responded at most to five items, note that these students have low success rate for this science subtest. Second column represents total number of responses which was calculated, for example, by multiplying total number of knowledge items (*n*=4) with total number of responses given to knowledge items (*n*=2206). Third column indicates the number of true responses given in each cognitive level. 65.23 % of total students correctly responded to 4.5% of total knowledge items, 2.26 % of total comprehension items, 2.25 % of total application items and 2.93 % of total analysis items. It is observed that the success level of students is 4.50 % for knowledge level. It is evident that these students have capability to answer knowledge level items better than other levels.

Table 12
The cognitive levels of students whose success percentile are 65.23 % (0-5)

The cognitive levels	Total number of responses	Number of true responses	Success rate%
Knowledge	8824	397	4.50
Comprehension	22060	499	2.26
Application	24266	547	2.25
Analysis	11030	323	2.93

Table 13 shows the success rate in each cognitive level for the students who correctly responded to six to ten items. According to Table 13, these students tend to answer the items in knowledge level correctly. 36.02% of the knowledge level item responses given by the students in this group are true. These students answered comprehension level items (22.20%) correctly the least. Their success in application (24.47%) and analysis level (26.17%) are comparable.

Table 13
The cognitive level of students whose success percentile are 16.97% (6-10)

The cognitive levels	Total number of responses	Number of true responses (n)	Success rate%
Knowledge	2296	827	36.02
Comprehension	5740	1274	22.20
Application	6314	1545	24.47
Analysis	2870	751	26.17

As seen in the Table 14, 6.89% of the students accurately answered 11-15 items. 52.79 % of this group answered correctly to in knowledge level and 38.67% in comprehension level, which shows that there is a sharp decrease in the success rate of comprehension level items. As stated Table 14, the success rate of application (41.28%) and analysis levels (43.43%) are approximately the same.

Table 14
The cognitive level of students whose success percentile are 6.89 % (11-15)

The cognitive levels	Total number of responses	Number of true responses (n)	Success rate%
Knowledge	932	492	52.79
Comprehension	2330	901	38.67
Application	2563	1058	41.28
Analysis	1165	506	43.43

Table 15 states the success rate (3.90%) of the students who correctly responded to 16-20 items out of 30 items. As observed from the Table 15, the success rates of these students are inversely proportional to the cognitive levels of the items with regards to Bloom’s Taxonomy. The knowledge level items (65.15%) are answered correctly the most compared to; comprehension (60.98 %), application (56.75%) and analysis levels (58.18%).

Table 15
The cognitive levels of students whose success percentile are 3.90% (16-20)

The cognitive levels	Total number of responses	Number of true responses (n)	Success rate%
Knowledge	528	344	65.15
Comprehension	1320	805	60.98
Application	1452	824	56.75
Analysis	660	384	58.18

In the Table 16, the success rate of students who responded 21-25 items correctly are shown. It is evident that the least success rate is in analysis level (71.97 %). These students have high success rate in comprehension level (80.36%). Among this group of students, the success rates from largest to least are comprehension (80.36%) level, knowledge level (79.56 %), application level (75.98%) and analysis level (71.97%).

Table 16
The cognitive levels of students whose success percentile are 4.05% (21-25)

The cognitive levels	Total number of responses	Number of true responses (n)	Success rate%
Knowledge	548	436	79.56
Comprehension	1370	1101	80.36
Application	1507	1145	75.98
Analysis	685	493	71.97

Table 17 represents the success rate of students who responded correctly to 26 to 30 items. As seen from Table 17, the least success rate is in analysis level (89.20%), which is expected. Students are most successful in knowledge level items (93.50%).

The next best success rate is in application level (93.18%), which is very close to knowledge level's rate. These rates are followed by the comprehension (92.10%) and analysis level (89.2%) rates.

Table17
The cognitive levels of students whose success percentile are 2.96% (26-30)

The cognitive levels	Total number of responses	Number of true responses (n)	Success rate%
Knowledge	400	374	93.50
Comprehension	1000	921	92.10
Application	1100	1025	93.18
Analysis	500	446	89.20

Summary

This chapter provided the detail information about the cognitive levels of the science subtest items, the difficulty parameters of the items, the cognitive levels of the students and the overall results of the analysis of this study. The answers to the research questions were given and supported with tables and figures. Most of the items are identified to be in low cognitive levels of Bloom's Taxonomy. Students encountered the most difficulty in science items when taking the 2007 UPE. We also observe that the numbers of the true answers given in application and analysis levels will increase as the total number of true answers given by students increases.

CHAPTER 5: DISCUSSION

Introduction

In this chapter, we mainly discuss the results of the study provided in Chapter 4. First part of this chapter consists of the overview of the study, major findings with regard to research questions and the conclusions. Second part is on implications for practice, implications for further research and limitations of the study.

Overview of the study

UPE have been carried out in our country for many years in order to place students to universities. The main purpose of this exam is to place students in suitable departments, based on the sufficiency of the students and criteria set by the universities.

This study emphasizes the importance of well-developed items in terms of Bloom's taxonomy by evaluating the success of the students. In this study, 2007 UPE science items were arranged according to the domain of cognitive levels of Bloom's Taxonomy. Bloom's Taxonomy is used because it provides a classification system to assess the cognitive skills of the students. The purpose of the arrangement of the science subtest is to examine the students' cognitive levels who took the corresponding exam. Students' responses in 2007 UPE science subtest are analysed. The analyses of the responses reveal both the cognitive levels of the students and the difficulty levels of the items.

Major findings

This study focused on the difficulty and cognitive levels of UPE science items, the cognitive levels of the students who took the corresponding exam and students' cognitive levels according to their answers to the science items. In the following sections, the major findings and possible reasons for these findings are discussed.

The cognitive levels of the items

According to the results of the analysis made by four pre-service teachers, the items are observed to be mostly in application level (36.67%), and then in order of frequency of appearance in the test they are in comprehension level (33.33%), analysis level (16.67%) and knowledge level (13.33%). This analysis proves that science items are usually designed to measure low cognitive levels of Bloom's Taxonomy.

According to researchers in this field, the number of the synthesis and the evaluation level items should make up a small percentage of the subtest (Köğçe 2005; Özmen, 2005). However, we observe that there are no items in synthesis and evaluation levels in the subtest used for this study. But this is expected for multiple-choice tests, which are not sufficient means to determine the high cognitive levels of the students with respect to Bloom's Taxonomy.

Findings of this study are analagous to those in literature. Many studies have shown that the items asked in UPE consist of mainly low thinking level items, which consist of knowledge, comprehension, and application level (Azar, 2005; Sesli Topçu, 2007). Kadayıfçı (2007) stated that the distribution of UPE chemistry item levels were determined to be in knowledge level (5.24%) the least, and average in comprehension (50.0%) and application levels (44.76%). It was determined that there was no higher order thinking level items such as analysis, synthesis and evaluation. Çevik (2010) classified 192 physics items that were asked between 2000-2008 UPE according to cognitive levels of Bloom's Taxonomy. Her study showed that 57% of the total physics items were in low cognitive levels (knowledge, comprehension and application), 43% of the total physics items were in higher order thinking level (analysis, synthesis and evaluation). The distribution of cognitive levels of items was balanced and expected.

In this study, our analysis on classification of the science items according to Bloom's Taxonomy supports the analysis in the literature. Most of the physics items (61.54%) were in application level, most of the chemistry items were in comprehension level (55.56%), and most of the biology items were in knowledge (37.50%) and comprehension level (37.50%) in 2007 UPE.

The UPE preparation test-books and written exams given at schools are divided into cognitive levels as stated in some of the research articles (Akpinar & Ergin, 2006; Azar, 2005; Kadayıfçı, 2007; Karamustafaoğlu et al. 2003). These studies show that students prepare for the university education with low cognitive level materials in the education system due to the anticipation of the level of items that is generally asked in the UPE. Hence, it can be deduced that schools are mostly focusing on training students through multiple-choice items. This creates a learning environment, where students are discouraged to explain their thought process through open-ended items and discussions, and they are not prepared to think analytically and critically in the universities.

Çoban and Hançer (2006) studied the UPE physics items asked between 1999-2003 and found that there is not a homogenous distribution among the cognitive levels of the items.

Furthermore, the frequent alterations in the examination system and the lack of standard measurement tools for these exams had caused inaccurate measurements and decline in the desirable teaching objectives in schools over the years. As a result of this system, schools and teachers have to adapt to those techniques and focus on the expected level items instead of teaching students to think broadly and critically. Consequently, the items, which are not certain in terms of measurement of thinking skills, are used for the purposes of important decisions in the UPE. Since students are only prepared for multiple-choice items, they do not develop writing skills and cannot express themselves professionally and scientifically (Berberoğlu, 2012).

The purpose of the UPE is to place undergraduate students, who have higher chance of being successful than other students, considering the quota of higher education programs. UPE should measure students' inquiry abilities and questioning skills; and also their knowledge and skills acquired in secondary education (Baykul, 1989).

The difficulties of the items

Before carrying out this study, we were expecting the lower cognitive level items to be correctly answered by more students, while higher cognitive levels items to be correctly responded by a small number of students. However, the results of the study did not exactly meet our expectations.

As this study has shown, items in knowledge level were expected to be easy for students. According to our analysis, 50% of knowledge level items were perceived as easy by the students, 50% of items were perceived as difficult. In comprehension level, approximately half of the students correctly answered some items in general. In contrast to these results, more than half of the students were expected to answer the comprehension level items (see Table 8.).

On the other hand, students could not answer low cognitive level items as expected. Although the items are in lower cognitive levels of Bloom's Taxonomy, the items may not measure the basic knowledge of the students. According to Berberoğlu (2013), studying habits and learning styles of the students might change throughout the preparation for multiple choice exams. Although these effects may not be observed in a short period of time, it can cause students, who spend a lot of time to solve multiple choice tests, to lose their problem solving skills, and their ability to think critically and analytically.

Research in the literature on cognitive levels of the UPE items supports that the items asked in UPE are overlapping with the items asked in written exams and textbooks (Morgil & Bayarı 1996). Even though all the students are thought the same curriculum in their respective schools, there can be differences across the regions, even schools in the same area in terms of emphasis. For example, some teachers might emphasize the topic of cell in biology, while the others might take cell topic in stride. Therefore, the emphasis on the curriculum will be affecting the true responses given to the items.

In this study, as expected students have difficulty at more than half of the application level items. The highest cognitive level of this test is analysis level. So, these items were expected to be answered by low percentage of students. According to this assumption, more than half of the analysis items were difficult for the students. However, two out of five analysis items were correctly answered by approximately half of the students.

According to the results mentioned above, the difficulty levels of the items were not balanced and proportional according to domain of cognitive levels of Bloom's Taxonomy. However, there is not any item, which is answered by more than 71 students out of 100. This shows that students often have difficulty in solving science items regardless of the cognitive level of the items. The reason might be that the items in the textbooks, which students use to prepare for the UPE, and written exam questions in schools are in lower cognitive levels of Bloom's Taxonomy or easy for students. The items in school exams and textbooks should include more analysis, synthesis and evaluation items, which would allow students to improve their understanding of the material as well as science literacy.

The available evidence seems to indicate that the cognitive level of biology items asked in the UPE remain in the lower levels of Bloom's taxonomy. According to Efe and Temelli (2003) 50% of the 36 items asked between 1999 and 2003 were seemed to be in medium difficulty. It was determined what kinds of skills students need to have in order to answer these items. The questions in this group are mostly in knowledge and comprehension, which require only basic biology knowledge. There was not any easy item in 1999 and 2001, but there were three items that was considered easy in 2000. Students had difficulty to solve 18 items in those three years. According to the results of the study, the cognitive levels of the items asked in

2000 are close to the cognitive levels of the students (Efe & Temelli, 2003). However, the cognitive levels of the items are higher than the cognitive levels of the students in 1999, especially in 2001 students had difficulty in solving the items, so it is seen that the success rate of the students was low.

The science subtest results of the 2015 UPE also shows that students have difficulty particularly in science items. The means of the science subtest were 6.48 out of 30 physics items, 8.75 out of 30 chemistry items, and 9.78 out of 30 biology items (MSPC, 2015). In 2016, the means of the science subtests were 5.03 out of 30 in physics, 9.53 out of 30 in chemistry, 7.73 out of 30 in biology (MSPC, 2016).

The reasons why students had difficulty in solving these items might be insufficient preparation for the fundamentals on the subject materials, the lack of explanation or selecting incorrect choice due to distraction. The preparation for the test items according to test techniques may cause students to overlook the main point of the material and instead rely on memorization. Hence, the cognitive levels of the items and the science scores of the students are not directly related.

Morgil and Bayarı (1996) points out that, even though the UPE physics items depend on basic physics knowledge and are in lower levels of Bloom's taxonomy; the materials for most of the items are given in the first two years of high school education. Hence, students, who do not master their skills on these topics in the first two years of high school, may have difficulty in solving these items.

The cognitive levels of the students according to their responses

When Table 13, Table 14, Table 15, Table 16, Table 17 and Table 18 were examined, it is obviously seen that the knowledge items were correctly responded mostly, which is expected. The ratio of responses in comprehension, application and analysis items vary among these tables.

The true answers for analysis items are relatively increasing with the increment in the true number of answers. However, the true answers for analysis items are still less than the knowledge level items. It shows that even if the number of the items increases that are correctly answered by students, they are more likely to answer knowledge level items correctly.

Responding to knowledge level items more than any other level shows that students are insufficiently equipped for higher levels of Bloom's Taxonomy. And students, who answer low cognitive level items more than high cognitive level items, may not use information to create something new, apply their prior knowledge on practice, deeply analyse a given topic, criticize and synthesize new information.

According to the results mentioned above, students have difficulty in solving high thinking level items more than low level items according to Bloom's Taxonomy, which is expected. Even if students have correctly answered high level of items, they appear to be susceptible to answer more low level of items. This study has shown that UPE items are not prepared to assess students' higher order thinking skills.

Implications for practice

In this context, UPE should not be a multiple-choice examination, because it is very challenging to prepare higher thinking level items in multiple-choice exams. It has been observed that when students graduate from high school they can do scientific research, identify problems, predict solutions, determine appropriate methods to solve problems, classify data collected from experiments and observations by creating tables, graphs, using statistical methods, analyse data, deduce results based on evidence, report and present scientific ideas (Secondary biology curriculum, 2013). It is very compelling to assess whether or not students gain these kind of skills stated above while studying for multiple choice test format.

Berberoğlu (2012) stated that the UPE should not be applied only multiple-choice exam format. Especially, the use of open-ended questions would have positive contribution to the quality of the education system. This study strongly supports the idea that the UPE should contain open-ended items that measure higher cognitive levels of the students, and it should include only a small amount of lower cognitive level items.

Inclusion of open-ended and high-level thinking items in the UPE would cause a modification on the textbook materials and school exams which would also be forced to include higher order thinking items. If the type of the examination would be changed from multiple choice to open ended, textbooks and other materials for the preparation to the exam would be changed accordingly. Therefore, students would be selected by a more accurate assessment to the departments of the universities that look for the specific qualifications such as being able to make analysis, thinking critically and analytically.

Moreover, this study shows that although UPE items are in low cognitive levels in Bloom's Taxonomy, students have difficulty in solving science items. Students may get enough UPE score to be placed at a university by simply responding to low level items. To avoid such situations, universities should determine their own criteria, have interviews and prepare their own exams for each department. Thus, students can be placed into universities according to the specifications set by the departments.

Implications for further research

- This study analysed the science subtest in UPE. Further studies may focus on different subject areas and the UPEs from different years.
- In this study, the classification of items was done based on Bloom's taxonomy. In further research, different taxonomies may be studied.
- In this study, quantitative descriptive design was used. Other designs such as causal-comparative designs can provide more information.
- Further research may be conducted using qualitative data by carrying out survey and interview including students' opinions to support the results.
- Further research can be done involving bigger data sets with different subgroups such as school types.

Limitations

The achievement rates of the students and cognitive levels of the items asked in UPE may differ year by year and according to the subject area. Only 2007 UPE science items and students' science score in corresponding year were analysed in this study. Excluding the data involving other subtests may have a slight effect on the results of the study. In this study, only a small number of students were used as sample. Including more students may produce more accurate results.

REFERENCES

Akpınar, E. & Ergin, Ö. (2006). Fen bilgisi öğretmenlerinin yazılı sınav sorularının değerlendirilmesi. *Milli Eğitim Dergisi, 172*, 225-231.

Alici, D., Başol, G., Çakan M., Kan A., Karaca, E., Özbek, Ö. Y., &Yaşar, M. (2011). *Eğitimde ölçme ve değerlendirme.* S.Tekindal (Ed.). Ankara: Pegem Akademi

Amer, A. (2006) Reflections on Bloom's revised taxonomy. *Electronic Journal of Research in Educational Psychology, 8,* 213-230

Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., & Wittrock, M.C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives.* New York: Longman.

Anderson, L. &Kratwohl, D. (2014). *Öğrenme öğretim ve değerlendirme ile ilgili bir sınıflama*. Özçelik, *D.* (Ed.). Ankara: Pegem Akademi.

Arı, A. (2013). Bilişsel alan sınıflamasında yenilenmiş bloom, solo, fink, dettmer taksonomileri ve uluslararası alanda tanınma durumları. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, *6*(2), 259-290.

Arslan, M. (2008). Günümüzde Montessori pedagojisi. *Milli Eğitim Dergisi*, *177*, 65-78.

Atılgan, H. (2006). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.

Ayiro, L. P. (2012). *A functional approach to educational research methods and statistics: qualitative, quantitative, and mixed methods approaches*. Lewiston, N.Y.: Edwin Mellen Press.

Azar, A. (2005). Analysis of Turkish high-school physics-examination questions and university entrance exam questions according to bloom's taxonomy. *Türk Fen Eğitim Dergisi, 2*(2), 68-74.

Bacanlı, H. (2005). *Gelişim ve öğrenme* (11th ed.). Ankara: Nobel Yayın Dağıtım.

Bakırcı, H. & Erdemir, N. (2010). Fizik öğretmen adaylarının mekanik konularını Bloom Taksonomisine göre öğrenebilme düzeyleri. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi, 3*(38), 81-91.

Baykul, Y. (1989) Öss ile yoklanan bilgi ve beceriler farklı okul tür ve sınıflarında ne ölçüde kazanılmaktadır, ÖSYM. Ankara: TAB Eğitim Yayınları.

Berberoğlu, G. (2012). Üniversiteye giriş nasıl olmalıdır? *Cito Eğitim: Kuram ve Uygulama, 15,* 9-21.

Berberoğlu, G., & İş Güzel, Ç. (2013). Eğitim sistemimizdeki ölçme ve değerlendirme uygulamaları nasıl olmalıdır? *Cito Eğitim: Kuram ve Uygulama, 21,* 9-16.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives; the classification of educational goals.*

Handbook I: Cognitive domain (pp. 2). New York, NY: David McKay Company.

British Columbia. (2013). *Large-scale assessment.* Retrieved from http://www2.gov.bc.ca/gov/topic.page?id=37F78E40586748EFBB2BDFAAB4D1AA4B

Burns, M. K. (2008). What is formative evaluation? *Minnesota Center for Reading Research.* Retrieved from http://www.cehd.umn.edu/reading/documents/faq/formativeeval.pdf

Burton, N. W. & Ramist, L. (2001). Predicting success in College: SAT studies of classes graduating since 1980. College Board Research Report no. 2001-2. New York: College Entrance Examination Board. Retrieved from http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2001-2-predicting-college-success-sat-studies.pdf

Bümen, N. T. (2006). Program geliştirmede bir dönüm noktası: Yenilenmiş Bloom taksonomisi. *Eğitim ve Bilim, 31*(142), 3-14.

Cambridge International. (2016). *AS & A Levels.* Retrieved from http://www.cie.org.uk/programmes-and-qualifications/cambridge-advanced/cambridge-international-as-and-a-levels/

Clarke, S. (2008). *Active learning through formative assessment.* London: Hodder Education an Hachette UK Company

Çakan, M. (2003). The role and importance of large-scale achievement. *Education and Science, 28*(128), 19-26.

Çetinsaya, G. (2014). *Büyüme, kalite, uluslararasilaşma: Türkiye yükseköğretimi için bir yol haritasi*. Eskişehir: Yükseköğretim Kurulu

Çevik, Ş. (2010). *Ortaöğretim 9,10 ve 11. sınıf fizik ders kitaplarında bulunan fizik soruları ile 2000-2008 yıllarında ÖSS'de sorulan fizik sorularının bloom taksonomisi açısından incelenmesi ve karşılaştırılması* (Unpublished dissertation). Dicle Üniversitesi Fen Bilimleri Enstitüsü, Fizik Anabilim Dalı, Diyarbakır.

Çoban, A. & Hançer, A.H. (2006). Fizik dersinin lise programları ve öss soruları açısından değerlendirilmesi. *Kastamonu Eğitim Dergisi*, *14*(2), 431-440.

Dave, R.H. (1970). *Psychomotor levels in developing and writing behavioural objectives.* R.J. Armstrong (Ed.). Tucson, Arizona: Educational Innovators Press.

DePascale A. C. (2003). The ideal role of large-scale testing in a comprehensive assessment system. *Journal of Applied Testing Technology, 5*(1), 1-11.

Durak H. İ. (2002) Eğitimde ölçme ve değerlendirmenin (sınama) öğretim üyeleri tarafından bilinmesi gereken temel ilkeleri. *Tıp Eğitimi Dünyası*, *7*, 43-46.

Efe, N. & Temelli, A. (2003). 1999-2000-2001 ÖSS biyoloji sorularinin düzey ve içerik yönünden değerlendirilmesi. *Kastamonu Eğitim Dergisi 11*(1), 105-114.

Erman, E. (2008).*2003-2006 yılları arasında yapılan orta öğretim kurumlarına öğrenci seçme sınavında yer alan tarih bilimi sorularının Bloom taksonomisine göre değerlendirilmesi* (Unpublished dissertation). Gazi

Üniversitesi Eğitim Bilimleri Enstitüsü, Ortaöğretim Sosyal Alanlar Eğitimi Tarih Eğitimi Bilim Dalı, Ankara.

Ertürk, S. (1972). *Eğitimde "program" geliştirme.* Ankara: MeteksanYayınları.

Finland Ministry of Education. (n.d.) *The finnish matriculation examination.* Retrieved from https://www.ylioppilastutkinto.fi/fi/english

Forehand, M. (2005). Bloom's taxonomy: Orginal and revised. In *Emerging Persceptives on Learning, Teaching, and Technology (e-Book)*. Retrieved 11 April, 2015 from http://eit.tamu.edu/JJ/DE/BloomsTaxonomy.pdf

Gruber, N. & Zdrahal Urbanek, J. (2006). *The Austrian education system - a short introduction.* Retrieved from http://www.ibw.at/html/bildungssystem/englisch06.pdf

Güneş, F. (2012). Skills and competencies set forth by bologna process in higher education. *Journal of Higher Education and Science, 2*(1), 1-9.

Gwet, K. (2001). *Handbook of inter-rater reliability*. Gaithersburg: StatAxis Publishing. (ISBN 0-9708062-0-5)

Head council of education and morality. (2013). *Secondary school biology (9-12) curriculum.* Ankara: Ministry of education.

Hesapçıoğlu, M. (1994). *Öğretim ilke ve yöntemleri* (3^{th} ed.) İstanbul: Beta Yayınevi

Harrow, A. (1972) *A taxonomy of psychomotor domain: A guide for developing behavioral objectives.* New York: David McKay

Huitt, W. (2011). Bloom et al.'s taxonomy of the cognitive domain. E*ducational psychology interactive*. Valdosta, GA: Valdosta State University. Retrieved from http://www.edpsycinteractive.org/topics/cognition/bloom.html [pdf]

Huitt, W. & Hummel, J. (2003). P*iaget's theory of cognitive development. Educational psychology interactive.* Valdosta, GA: Valdosta State University. Retrieved from http://www.newriver.edu/images/stories/library/Stennett_Psychology_Articles/Piagets%20Theory%20of%20Cognitive%20Development.pdf

IBO. (2016). *Assessment and exams.* Retrieved from http://www.ibo.org/programmes/diploma-programme/assessment-and-exams/

IGCSE. (2016). *Cambridge IGCSE curriculum.* Retrieved from http://www.cie.org.uk/programmes-and-qualifications/cambridge-secondary-2/cambridge-igcse/curriculum/

Kadayıfçı, K.G. (2007). *Liselerde ve öss sorularında sorulan kimya sorularınınprograma uygunluğunun incelenmesi* (Unpublished dissertation). Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ortaöğretim Fen ve Matematik Alanları Eğitimi Ana Bilim Dalı, Ankara.

Kan, A. (2006) *Eğitimde ölçme ve değerlendirme***.** Ankara: Ertem Matbaası.

Karamustafaoğlu, S., Sevim, S., Karamustafaoğlu O. & Çepni, S. (2003). Analysis of Turkish high-school chemistry-examination questions according to Bloom's taxonomy. *Chemistry Education: Research and Practice, 4*(1), 25-30.

Kutlu, Ö. (2003). Cumhuriyetin 80. Yılında: ölçme ve değerlendirme. *Milli Eğitim Dergisi, 160.* Retrieved from http://dhgm.meb.gov.tr/yayimlar/dergiler/Milli_Egitim_Dergisi/160/kutlu.htm

Küçükahmet, L. (2005). *Öğretimde planlama ve değerlendirme* (17th ed.). Ankara: Nobel Yayın Dağıtım.

Köğçe, D. (2005). *Öss sınavı matematik soruları ile liselerde sorulan yazılı sınavsorularının bloom taksonomisine göre karşılaştırılması* (Unpublished dissertation). Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İlköğretim AnabilimDalı.

Köğçe, D., & Baki, A. (2009). Matematik öğretmenlerinin yazılı sınav soruları ile ÖSS sınavlarında sorulan matematik sorularının Bloom taksonomisine göre karşılaştırılması. *Pamukkale Universitesi Eğitim Fakultesi Dergisi, 26,* 70-80.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: an overview. *Theory into Practice, 41*(4), 212-218.

Leahy, S., Lyon, C., Thompson M. & Wiliam D. (2005). Classroom assessment: minute by minute, day by day. *Assessment to Promote Learning, 63*(3). 19-24.

Loveless, A. & Ellis, V. (2002). Information and communication technologies, pedagogy and the curriculum. *Education and Information Technologies, 7*(1) 81-83.

MoNE (Ministry of Education). (2012). *12 years compulsory education*. Retrieved from http://www.meb.gov.tr/duyurular/duyurular2012/12yil_soru_cevaplar.pdf

Morgil, F. İ. & Bayarı, S. (1996). Öss ve ÖYS fizik sorularının soru alanlarına göre dağılımı, çözülebilirlikleri ve başarının bağlı olduğu etkenler. *Hacettepe Üniversitesi Eğitim Fakültesi, 12,* 215-220.

MSPC. (2006). *Öğrenci seçme sınavı (ÖSS) kılavuzu.* Ankara: Yüksek Öğretim Kurumu, Öğrenci Seçme ve Yerleştirme Merkezi. Retrieved from http://osym.gov.tr/belge/1-6450/2006-osys-kilavuzu.html

MSPC. (2007). *Öğrenci seçme sınavı (ÖSS) soru kitapçığı.* Ankara: Yüksek Öğretim Kurumu, Öğrenci Seçme ve Yerleştirme Merkezi. Retrieved from http://dokuman.osym.gov.tr/pdfdokuman/arsiv/2007OSS/oss_genel_aciklama.pdf

MSPC. (2007). *2007- ÖSS'ye ilişkin sayısal bilgiler.* Ankara: Yüksek Öğretim Kurumu, Öğrenci Seçme ve Yerleştirme Merkezi. Retrieved fromhttp://osym.gov.tr/belge/1-8314/sayisal-bilgiler.html

MSPC. (2008). *TC Yükseköğretim kurumu öğrenci seçme ve yerleştirme merkezi 2008 yılı faaliyet raporu.* Ankara: ÖSYM

MSPC. (2009). *2009 ÖSYS kılavuzu.* Ankara: ÖSYM.

MSPC. (2010). *2010-ÖSYS Yükseköğretim programları ve kontenjanları kılavuzu.* Ankara: ÖSYM

MSPC. (2012). *TC Yükseköğretim kurumu öğrenci seçme ve yerleştirme merkezi 2012 yılı faaliyet raporu.* Ankara: ÖSYM

MSPC. (2014). *Tarihsel gelişme.* Retrieved from http://www.ÖSYM.gov.tr/belge/1-2706/tarihsel-gelisme.html

MSPC. (2015). *2015-LYS sonuçlarına ilişkin sayısal bilgiler*. Retrieved from http://dokuman.osym.gov.tr/pdfdokuman/2015/LYS/2015LYSSAYISALBILGILER30062015.pdf

MSPC. (2015). *2015-YGS sonuçlarına ilişkin sayısal bilgiler*. Retrieved from http://dokuman.osym.gov.tr/pdfdokuman/2015/YGS/2015-YGSSAYISALBILGILER19032015.pdf

MSPC. (2016). *2016-LYS sonuçlarına ilişkin sayısal bilgiler*. Retrieved from http://dokuman.osym.gov.tr/pdfdokuman/2016/LYS/LYSSayisalBilgiler19072016.pdf

Neuman, W. L. (2000). *Social research methods: qualitative and quantitative approaches* (7th ed). Pearson. Available from http://www.learningace.com/textbooks/12346-social-research-methods-qualitative-and-quantitative-approaches-7th-edition

NUFFIC. (2016). Higher education in the Netherlands. Retrieved from https://www.studyinholland.nl/documentation

Online Curriculum Centre. (2014). *The diploma programme.* Retrieved from http://occ.ibo.org/ibis/occ/guest/dpyyy_home.cfm?subject=dpyyy&CFID=72

5675&CFTOKEN=73866983&jsessionid=bc30e20840fb95f0d6f05b5344c671cad46d

OECD. (2010). *PISA 2009 Results: What students know and can do – student performance in reading, mathematics and science (Volume I).* Retrieved from http://www.oecd-ilibrary.org/education/pisa-2009-results-what-students-know-and-can-do_9789264091450-en

Özcan, S. & Oluk, S. (2007). Analysis of questions used in science lessons at primary school according to Piaget and Bloom Taxonomy. *D.Ü. Ziya Gökalp Eğitim Fakültesi Dergisi, 8,* 61-68.

Özçelik, D. (1997). *Ölçme ve degerlendirme.* Ankara: ÖSYM Yayınları.

Özmen, H. (2005). 1990–2005 ÖSS sınavlarındaki kimya sorularının konu alanlarına ve Bloom taksonomisi'ne göre incelenmesi. *Eurisian Journal of Educational Research, 21*, 187-199.

Özmen, H. & Karamustafaoğlu, O. (2006). Lise II. sınıf fizik-kimya sınav sorularının ve öğrencilerin enerji konusundaki başarılarının bilişsel gelişim seviyelerine göre analizi. *Kastamonu Eğitim Dergisi*, *14*(1). 91-100.

Piaget, J. (1964). Cognitive development in children: Piaget. *Journal of Research in Science Teaching, 2,* 176-186. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/tea.3660020306/pdf

PISA. (2016). *PISA Türkiye, PISA nedir?* (22 June 2016) Retrieved from http://pisa.meb.gov.tr/?page_id=18

Sadler, D. R. (1998). Formative assessment: Revisiting the territory, *Assessment in Education: Principles, Policy & Practice, 5*(1), 77-84.

Sesli Topçu, A. (2007). Biyoloji öğretmenlerinin yazılı sınav soruları ile ÖSS sorularının bloom taksonomisine göre karşılaştırmalı analizi. Retrieved from YÖK Tez Merkezi. (AAT 212083)

Shewhart, W. A. (1931). Random sampling. *The American Mathematical Monthly, 38*(5), 245-270.

Simpson, E.J. (1972). *The classification of educational objectives in the psychomotor domain.* Washington, DC: Gryphon House.

Sönmez, V. (1997). *Sosyal bilgiler ögretimi ve ögretmen kılavuzu.* Ankara: Anı Yayıncılık.

Sönmez, Ö. F., Koç, H. & Çiftçi, T. (2013). ÖSS, YGS ve LYS sınavlarındaki coğrafya sorularının Bloom taksonomisi bilişsel alan düzeyi açısından analizi. *Karadeniz Araştırmaları Dergisi, 36*, 257-275.

Stebbins, R. A. (2001). *Exploratory research in the social sciences. Sage university paper series on qualitative research methods* (pp. 9). Thousand Oaks, CA: Sage.

Struwig, F. W. & Stead, G. B. (2001). *Planning, designing and reporting research.* South Africa: Pearson.

Taras, M. (2005) Assessment – summative and formative – some theoretical reflections, *British Journal of Educational Studies, 53*(8), 466-478.

TED. (2005). *Türkiye'de üniversiteye giriş sistemi araştırması-Sonuç raporu.* Ankara: Türk Eğitim Derneği.

Tezbaşaran, A. A. (1994). ÖSYS testlerinde yoklanmak istenen bilişsel davranışlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 10,* 79-84.

TIMSS. (2015). *Tanıtım kitapçığı.* (22 June 2016) Retrieved from http://timss.meb.gov.tr/wp-content/uploads/Tanitim_Kitapcigi.pdf

Turgut, M.F. (1977) *Eğitimde ölçme ve değerlendirme metotları*. Ankara: Saydam Matbaacılık

Turkish council of higher education. (2007). *Türkiye'nin yükseköğretim stratejisi.* (pp, 13). Ankara: Yükseköğretim Kurulu. (ISBN: 978-975-7912-32-3)

Turkish Council of higher education. (2016). *Üniversitelerimiz.* Retrieved from http://www.yok.gov.tr/web/guest/universitelerimiz

Tutkun, Ö. F. (2012) Bloom'un yenilenmiş taksonomisi üzerine genel bir bakış. *Sakarya University Journal of Education*, *2*(1), 14-22.

UKCAT. (2016). *What is the UKCAT?* Retrieved from http://www.ukcat.ac.uk/about-the-test

Urbach, P. (1989) Random sampling and the principles of estimation. *Proceedings of the Aristotelian Society New Series*, *89,* 143-164.

Varış, F. (1996). *Eğitimde Program Geliştirme.* Ankara: Alkım Kitapçılık Yayıncılık.

Yurdabakan, İ. (2012). The effects of Bloom's revised taxonomy on measurement and evaluation in education. *Gaziantep Üniversitesi Sosyal Bilimler Dergisi, 11*(2), 327-348.