THE COMPARISON OF PISA AND TIMSS FOR FRAMEWORK AND ACHIEVEMENT

EMİNE ÖZGE YILMAZ

SEPTEMBER 2010

THE COMPARISON OF PISA AND TIMSS FOR FRAMEWORK AND ACHIEVEMENT

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

BY

EMİNE ÖZGE YILMAZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN THE DEPARTMENT OF
ELEMENTARY SCIENCE AND MATHEMATICS EDUCATION

SEPTEMBER 2010

Approval of the Graduate School of Social Sciences

____

Prof. Dr. Meliha ALTUNIŞIK
Director

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

____

Prof. Dr. Hamide ERTEPINAR
Head of Department

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

______________________	______________________
Prof. Dr. Giray BERBEROĞLU Co-Supervisor	Assoc. Prof. Dr. Özgül YILMAZ-TÜZÜN Supervisor

Examining Committee Members

Assoc. Prof. Dr. Özgür ERDUR BAKER	(METU, EDS)	______________________
Assoc. Prof. Dr. Özgül Yılmaz-Tüzün	(METU, ELE)	______________________
Prof. Dr. Giray Berberoğlu	(METU, SSME)	______________________
Assoc. Prof. Dr. Semra Sungur	(METU, ELE)	______________________
Assist. Prof. Dr. Elvan ŞAHİN	(METU, ELE)	______________________

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

Name, Last name: Emine Özge, YILMAZ

Signature:

ABSTRACT

THE COMPARISON OF PISA AND TIMSS FOR FRAMEWORK AND ACHIEVEMENT

Yılmaz, Emine Özge
M.S., Department of Elementary Science and Mathematics Education
Supervisor : Assoc. Prof. Dr. Özgül Yılmaz-Tüzün
Co-Supervisor: Prof. Dr. Giray Berberoğlu

September 2010, 98 pages

The present study aimed to (1) Investigate the comparability of PISA 2006 and TIMSS 2007 (2) Analyze the similarities and differences of frameworks of PISA 2006 and TIMSS 2007. In accordance with the purpose of the study, content analysis was done. Content analyses showed that content of two international studies were similar to each other in some extent. The competencies and cognitive skills were also similar on surface whereas the base of the competencies and cognitive skills were different. In order to find similarities and differences, the items assessing the similar cognitive skills were selected for PISA and TIMSS. The analysis of the items in terms of the percentage of Turkish students' correct answers demonstrated that while identifying scientific issues in PISA and knowing in TIMSS were not similar, explaining phenomena scientifically in PISA and applying in TIMSS and using scientific evidence in PISA and reasoning in TIMSS were similar.

Keywords: PISA 2006, TIMSS 2007, Achievement, Framework Comparison

ÖZ

PISA VE TIMSS'İN KAPSAM VE BAŞARI YÖNLERİNDEN
KARŞILAŞTIRILMASI

Yılmaz, Emine Özge
Yüksek Lisans, İlköğretim Fen ve Matematik Alanları Eğitimi Bölümü
Tez Yöneticisi : Doç. Dr. Özgül Yılmaz-Tüzün
Ortak Tez Yöneticisi: Prof. Dr. Giray Berberoğlu

Eylül 2010, 98 sayfa

Bu çalışmada (1) PISA 2006 ve TIMSS 2007'nin karşılaştırılabilirliğini analiz etmek (2) PISA 2006 ve TIMSS 2007'nin kapsamlarının benzerlik ve farklılıklarının analiz edilmesi amaçlanmıştır. Çalışmanın amacı doğrultusunda, İçerik Analizi kullanılmıştır. İçerik analizi, iki uluslararası çalışmanın içeriklerinin bir ölçüde benzer olduklarını göstermiştir. Yeterlikler ve bilişsel beceriler de görünüşte benzer olsa da temelleri farklıdır. Farklılık ve benzerlikleri bulmak için PISA ve TIMSS için benzer bilişsel becerileri ölçen sorular seçilmiştir. Türk öğrencilerin doğru cevap yüzdelerine göre soru analizleri PISA'daki bilimsel soruları tanımlama ve TIMSS'deki bilme için farklılık gösterirken, PISA'daki bilimsel olguları açıklama ve TIMSS'deki uygulama ile PISA'daki bilimsel delilleri kullanma ve TIMSS'deki akıl yürütmenin benzer olduğunu göstermiştir.

Anahtar Kelimeler: PISA 2006, TIMSS 2007, Başarı, Kapsam Karşılaştırması

To my family…

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Assoc. Prof. Dr. Özgül Yılmaz-Tüzün, the supervisor of my thesis, for her encouraging efforts and constructive criticism. I am also thankful to Prof. Dr. Giray Berberoğlu for his constructive criticisms about the study.

I would like to thank my family for their patience, endless love and moral support.

I would also like to thank my close friends for their invaluable moral support, encouragement, and patience throughout the process. I love you very much.

I thank the members of my committee for their willingness to serve on the committee and their valuable feedback.

Finally, I thank TÜBİTAK for their scholarship during my master studies.

Thank you all very much indeed.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURE

CHAPTER 1

INTRODUCTION

In recent years, international studies of educational achievement draw not only the public attention but also governments' and policy makers' attention. These international studies reveal the educational achievement differences among countries (e.g., Beaton et al., 1999). An increasing number of countries have participated in these large-scale international studies to reveal differences in terms of their students, classrooms, and school levels. In addition these studies also aimed to explain the potential reasons of these observed differences (Creemers, 2006).

Large-scale international studies provide valid and reliable cross-national information about positive and negative aspects of students' educational achievement among countries. For these studies the quality of data was high in general because of their clearly defined population, sampling process, well-developed frameworks and instruments including translation into the different languages, well-structured process of administration in all countries, and well-developed scoring of students' responses (Olsen, 2005). More specifically, in all countries, the population and sampling had clearly defined. Moreover, the frameworks of international studies were developed in detail and the translation process was well defined in technical reports of the studies. The process of administration was controlled by the appointed people and also defined in the guidebooks. The high degree of reliability and validity of the instruments and data collected through these international studies ensure comparability in terms of students' achievement among the countries. Turkey has taken part in two of these international studies including science; Programme for International Student Assessment (PISA) in the year of 2003 and 2006 and Trends in International Mathematics and Science Study (TIMSS) in the year of 1999 (TIMSS-R) and 2007. For the purpose of this study, the science frameworks of PISA and TIMSS were compared and they were further investigated to explore students' weaknesses and strengths in the major science areas.

1. 1 The Programme for International Student Assessment (PISA)

One of the large scale international studies is the Programme for International Student Assessment (PISA) which measures knowledge and skills of students who are aged between 15 years 3 months and 16 years 2 months. Students at this age in most OECD countries have ended their compulsory schooling of ten years of education (PISA Technical Report, 2009).

PISA 2006 contributed three following outcomes. The first one was the information about the profile of students' knowledge and skills. The second one was relating contextual indicators such as demographic, social, economic and educational variables to those skills. The third one was changes in student-level and school-level background variables and outcomes (PISA Assessment Framework, 2006).

PISA has been administered by Organization for Economic Cooperation and Development (OECD) in three year cycles in the areas of reading, science and mathematics. In PISA, some of the items were used repeatedly in consecutive applications because with the repeated data sets, it would be easier to assess trend over time in both performance of students in their countries and the relative performances of students across countries.

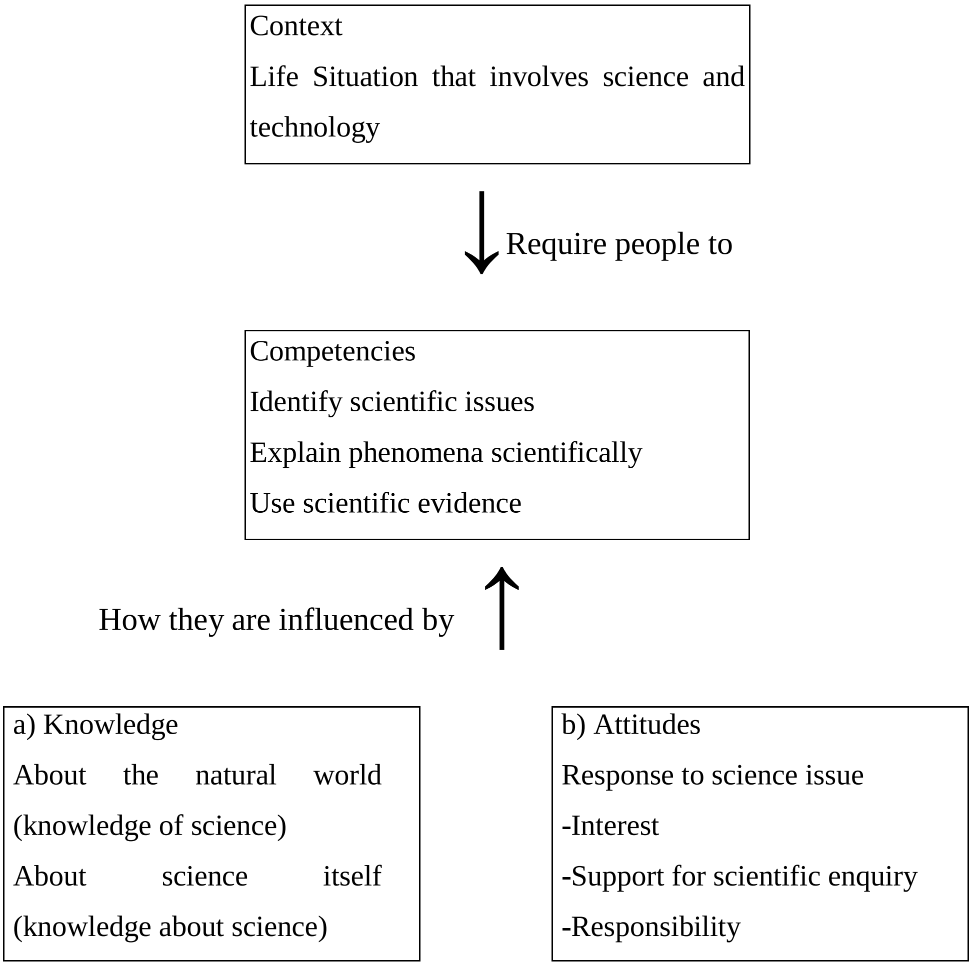
PISA was first conducted in 2000 with the focus on reading literacy but the assessment also included science and mathematics in addition to student, family and institutional factors affecting the performances of students in 43 countries. In 2003, the particular subject area was mathematical literacy measured across 41 countries including Turkey. Besides measuring student capabilities in reading, mathematics and science, cross disciplinary problem solving ability in mathematics and science were assessed in 2003. In the year of 2006, PISA was carried out among more than 400 000 students across 57 countries with all OECD countries, as well as 27 partner countries. Turkey was also a participant of PISA 2006 with 160 elementary and secondary schools and 4942 students (EARGED, 2007).

The purpose of PISA is to provide information on the following questions:

- "How well are young adults prepared to meet the challenges of the future?
- Are they able to analyze reason and communicate their ideas effectively?
- Do they have kinds of interests they can pursue throughout their lives as productive members of the economy and society (OECD PISA, 2010, para. 1)?

In light of above questions one can see that PISA measured not only the knowledge of the subjects but also application and interpretation of the knowledge in the daily life situation. PISA also emphasized on life-long learning by asking students about their motivation to learn, attitudes to learn and their beliefs about themselves. In addition to that PISA provided a successful adaptation to a changing world with acquiring new knowledge and necessary skills. Moreover, PISA had a different content than other international surveys. The content of PISA included the knowledge which students would use in their future life rather than the curricula. Therefore, it could be said that PISA measures students' preparedness for adult life at the end of the compulsory education (PISA Assessment Framework, 2006).

PISA 2006 developed its science assessment task and questions via the following four interrelated aspects as seen in Figure 1.1.



Note. From *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006* (p.26), 2006, Paris: OECD. Copyright 2006 by OECD. Reprinted with permission.

Figure 1.1 The Framework for PISA 2006 Science Assessment

PISA measured students' capacity in identifying scientific issues, explaining phenomena scientifically and using scientific evidence in their daily life with interpreting, solving, and making decisions related with science and technology via those four aspects as seen in Figure 1.

Although many studies such as TIMSS and PIRLS **(**Progress in International Reading Literacy Study) assessed students' knowledge of specific science content, PISA had a different approach towards science. PISA approach defined scientific competency as involving many aspects of life from success at work to active citizenship. According to OECD, students could not be successful in their future life only by memorizing just the scientific facts. They should analyze what is actually needed and find out solutions when there are no clear rule-based solutions (PISA, 2007). In PISA 2006, according to four areas given in Figure 1, countries have showed different achievement levels in major science areas and competencies stated in PISA 2006.

1.2 Trends in International Mathematics and Science Study (TIMSS)

TIMSS is another international large scale study about science and mathematics education for the fourth and eighth grade students. TIMSS, which is conducted in regular 4-year cycle, is a project of the International Association for the Evaluation of Educational Achievement (IEA). IEA is an independent international cooperation of national research institutions and government agencies.

TIMSS was first conducted in 1995 with 41 participant countries. In 1999, Turkey was one of the participants of 38 countries of the Third International Mathematics and Science Study-Repeat, referred to TIMSS-R. Then, in 2003 there were 48 countries participated at the eighth grade and 26 countries at the fourth grade but Turkey did not participate in both grades. In 2007, TIMSS was carried out in 59 countries around the world and Turkey was also one of the participants at eight grades.

The aim of TIMSS was to provide monitoring the implementation of educational policies and guiding educational decision making and better policies around the world. Furthermore, TIMSS gave comparative information about educational achievement with the aim of improving teaching and learning in mathematics and science. TIMSS also collected a rich array of information about gender differences in achievement, students' background, attitudes towards science, science curriculum, science teachers' background characteristics in addition to their education and training, classroom characteristics and instructions and school contexts for learning science. TIMSS provided comprehensive and internationally comparable data at the fourth and eighth grades to the countries. Therefore, participant countries would evaluate their mathematics and science teaching across time and across grade. TIMSS also helped countries in monitoring the progress in learning from the fourth grade to eighth grade with its cyclic design. As well as learning progress, TIMSS monitored the relative effectiveness of teaching comparably fourth grade with eighth grade. Participant countries also have inferences about the contexts which students learn best by the comparisons with other countries having higher achievement results (TIMSS, 2008).

Science Assessment Framework organized in three dimensions namely content domain, cognitive domain or thinking processes and contextual domain in TIMSS 2007. The details were given in the Table 1.1.

Table 1.1 The Science Framework for the TIMSS 2007

Content Domain	Cognitive Domain	Contextual Domain
Biology	Knowing	Curriculum
Chemistry	Applying	Schools
Physics	Reasoning	Teachers and their preparation
Earth Science		Classroom activities and characteristics

Curriculum was considered as a broad explanatory factor in explaining students' achievement (Robitaille & Garden, 1996). TIMSS used the curriculum mainly in discovering the students' access to educational opportunities and the factors affecting students' access to educational opportunities. There were three dimensions in the TIMSS curriculum model. These were the intended curriculum, the implemented curriculum and the achieved curriculum. The intended curriculum showed intentions of society for students what to be taught and how the educational organizations should be done in order to facilitate learning. The implemented curriculum was what was really taught in classrooms and who taught it and how it was taught. The achieved curriculum meant what students actually learn in classrooms. This modeling was used for explaining students' learning by using mathematics and science tests with questionnaires giving information about student characteristics (TIMSS Assessment Framework, 2005). As in PISA 2006, the performances of countries in science differed in the TIMSS 2007.

1.3 The Purpose of the Study: Investigating Comparability of Science Aspects of PISA and TIMMS

Not only Turkey, as shown on Table 1.2, more and more countries are taking part in these international studies since 1995. The ranking between countries can inform educationalists about the national strengths and weaknesses (Lie, 2005, Çalışkan et al., 2010). For most of the countries, the ranking, in other words the achievement in the international studies are important because the results from TIMSS and PISA provided future decisions on planning and implementation of educational policies resulting from achievement in the assessments (Olsen, 2004; Wu, 2006). One of the countries was Turkey which the Ministry of National Education has explained one of the needs for the new curriculum with the results from international studies such as TIMSS and PISA (Educational Board of Turkey, 2005).

Table 1.2 The Number of Countries Participating in PISA and TIMSS

PISA 2000	PISA 2003	PISA 2006	PISA 2009
43	41	57	65
TIMSS 1995	TIMSS 1999	TIMSS 2003	TIMSS 2007
41	35	49	60

Several studies (Acar, 2008; Beaton et al., 1999; Hutchison & Schagen, 2007; Olsen, 2004) claimed that there should be more secondary analysis on international studies. In the report of The Economic Policy Research Foundation (EPRF), Acar (2008) emphasized that the results of PISA should be investigated because the results both help finding the weaknesses of educational policies of countries and giving clues about the factors enabling good performance in PISA. Moreover, Acar declared that PISA results could be used as a success predictor in future life of students. Acar (2008) exemplified Canada and Denmark as a support. In the report of PISA 2000, only 30 % of students from lower achievers continued to university, while the number for high achievers continuing to a university was 90 % for Canada (Acar, 2008).

There are several studies comparing the achievement of students in PISA and TIMSS in different countries (Bybee, 2007; Hutchison & Schagen, 2007; Nohara, 2001; Ruddock et al. 2006) while there were studies claiming that there could not be a direct comparison between TIMSS and PISA (Adams, 2003; Harding, 2003; Olsen, 2004).

Taking into account achievement difference of USA students, Nohara (2001) studied the comparison of PISA 2000, TIMSS-R and NAEP 2000. In his study, the frameworks and purposes of three studies were compared. Bybee (2007) also analyzed PISA and TIMSS 2003 in a general way. Bybee compared the achievement of students in USA with other countries participating TIMSS and PISA. Another comparison between PISA and TIMSS 2003 in terms of aims, sampling and items by using IRT method was done by Hutchison and Schagen (2007). They have also researched the achievement difference of countries in TIMSS and PISA and made possible explanations about the difference. In another study, Ruddock et al. (2006) analyzed the appropriateness of items in PISA and TIMSS for the English students and explained the possible achievement difference.

Harding (2003) and Adams (2003) were the researchers saying there could not be direct comparison between PISA and TIMSS. The authors explained the reason as the different target population in PISA and TIMSS. Another researcher asserting that direct comparison wouldn't be done between PISA and TIMSS was Olsen (2004). The author stated that PISA and TIMSS differ in their assessing aims. Another reason for that explained in the study of Olsen (2005) as the difference of measured scientific concepts in two studies.

There were a limited number of studies searching specifically on competencies or cognitive skills (Ciascai, 2009; Kjærnsli, 2003; Kjærnsli & Lie, 2004).

Ciascai (2009) analyzed the similarities of Romanian curriculum and TIMSS framework via Bloom's Taxonomy. Kjærnsli (2003) investigated the achievement differences of Nordic countries in TIMSS 1995 by classifying science in two groups as scientific processes and demonstrating understanding of scientific concepts.

In another study, Kjærnsli and Lie (2004) explored the differences and similarities of countries in achievement in terms of gender and item analysis across Nordic countries in PISA 2000. Countries' achievement was investigated via two categories defined as scientific processes and scientific knowledge (Kjærnsli and Lie, 2004).

As seen from these researches, several countries were studying of achievement in international studies in terms of item analysis or comparison with other countries. In addition to that achievement comparison between PISA and TIMSS was also studied in terms of aims, sampling, and domains of test and item types in a general way. Although there were such studies comparing PISA and TIMSS, several studies asserting that there could not be direct comparisons between PISA and TIMSS.

In this study, the comparability of PISA and TIMSS in terms of Turkey was investigated.

Since the comparability of TIMSS and PISA were questioned, it was also important to investigate the similarities and differences of frameworks of PISA and TIMSS.

The research questions of this study were:

1) Are the PISA and TIMSS comparable?
2) Are the science frameworks of TIMSS and PISA parallel?

1.4 The Significance of the Study

In this study, the comparability of PISA and TIMSS were investigated. Although the names of major areas in science differed in PISA and TIMSS, there were several similarities on surface. The item analysis of PISA and TIMSS enabled the achievement of Turkish students in two studies. Therefore, international studies, with high reliability and validity, comparison of the science achievement in

TIMSS and PISA would give clues of weak areas in Turkish science curriculum. By using students' results from these international studies, the effectiveness of currently implemented Turkish science and technology curriculum may be addressed.

According to Hutchison and Schagen (2007), more analysis should be done on TIMSS and PISA because they gave the clues of how a national system would have a high quality education. Therefore, it is significant to explore the weaknesses and strengths of Turkish students in PISA and TIMSS. In the light of results from international studies, policy makers could change or adapt the curriculum to have a high quality of education.

One of the benefits of the international studies was to provide information about differences and similarities of institutions, systems and practices among countries (White & Smith, 2005). In addition to policy makers, teachers could also use the results of these international studies. For instance, teachers could use the weak and strong areas of Turkish students and arrange the lesson plans and teaching methods according to results of this study. Policy makers can investigate the trends in both PISA and TIMSS over years and compare the results with this study. Thus, policy makers can make changes in the lesson hours or in objectives of the lessons.

CHAPTER 2

LITERATURE

In the literature, the frameworks of the Programme for International Study (PISA) and the Trends in Mathematics and Science Studies (TIMSS) and the relevant research studies about the comparison of the frameworks were presented.

2.1 Science Assessment Framework in PISA

2.1.1 The Development Process for the PISA 2006 Assessment Framework

The development of PISA framework has been an ongoing process since 1997. At first the definition of the assessment domain and descriptions of the assumptions in those domains were determined. In the second step, the way to organize the tasks in the domains in order to explain student achievement was evaluated. The key characteristics of the domains were operationalized in the test construction process according to definitions in the existing literature and experience from other large-scale assessments. Then, in order to understand task difficulty across countries, the validation of the variables and assessment of the contribution were done. Finally, the scheme explaining the results was prepared (PISA, 2007).

The main advantage of preparing a framework was explained as improvement of assessment in PISA 2006. Also there were other advantages like creating a common language for explaining the purpose of the assessment and what the assessment desired to measure. This would be helpful in making a common decision around framework and define the assessment goals. Furthermore, the analysis on the successful performance in knowledge and skills provided a basis for establishing levels of proficiency. The key variables for a successful performance could be determined by a framework and so frameworks helped in making changes in the assessment over time. Moreover, the connection of the variables and student performance gave an important link between the public policy and assessment improving the usefulness of data (PISA, 2007).

The assessment in PISA focused on what 15 year old students should know and be able to do in order to get prepared for future life. The overall purpose of PISA science assessment could be best described with “Scientific Literacy” (PISA, 2006).

According to Turmo (2004), the term “Scientific Literacy” was first used in publication by Paul Hurd in the *Science Literacy: Its meaning for America Schools* (Turmo, 2004). Since then, there were many definitions of “Scientific Literacy” but in PISA 2006, Scientific Literacy was defined as in the following way.

> "An individual's scientific knowledge and use of that knowledge to identify questions, to acquire new knowledge, to explain scientific phenomena, and to draw evidence based conclusions about science-related issues, understanding of the characteristic features of science as a form of human knowledge and enquiry, awareness of how science and technology shape our material, intellectual, and cultural environments, and willingness to engage in science-related issues, and with the ideas of science, as a reflective citizen (PISA Assessment Framework, 2006, p. 12)."

In PISA 2006, the assessment of scientific literacy was very important. The definition gave a continuum from less scientifically literate to more scientifically literate. Rather than "science", scientific literacy emphasized the use of scientific knowledge in daily life. PISA 2006 aimed to assess both the cognitive and affective aspects of students' scientific literacy. In PISA 2006 that the areas used in scientific literacy were namely Context, Competencies, Knowledge and Attitudes. These areas were explained as follows.

2.1.1.1 Knowledge

In PISA 2006, science literacy included two types of knowledge, namely "Knowledge of science" involving to understand main scientific concepts and theories and "Knowledge about science" involving to understand the nature of science. Knowledge of science meant to understand basic scientific concepts and theories. In other words, knowledge of science referred to knowledge about the natural world. PISA gave importance to application of the major fields such as physics, chemistry, biology, Earth and space science and technology to real-life situations as knowledge of science. Therefore, assessment material had to be relevant to real-life situations, representative of important scientific concepts and appropriate to the developmental level of 15-year-olds. There were four content areas which were important in adults' understanding was the natural world in personal, social and global contexts. In spite of traditional naming such as "Physical Sciences" in the content, PISA represented a different name namely "System" for contents to show people the meaning of varied concepts and to emphasize the importance of the relationship via contents. In PISA, the four content names were "Physical Systems", "Living Systems", "Earth and Space Systems" and "Technology Systems" (PISA Assessment Framework, 2006).

On the contrary, knowledge about science meant to understand science as a human activity and also to understand both power and limitations of scientific knowledge. PISA defined knowledge about science in two categories namely as "Scientific Enquiry" and "Scientific Explanations". Scientific Enquiry was the main science procedures and several components of that procedure. The results of "scientific enquiry" are the "Scientific Explanations" (PISA, 2007).

2.1.1.2 Competencies

In PISA 2006, there were three competencies students should have in order to answer the science questions. In selecting the competencies, there were some criterias such as the importance in applying scientific knowledge, connection to the key cognitive abilities such as inductive/deductive reasoning, systems-based thinking, critical decision making, transformation of information construction and communication of arguments and explanations based on data, thinking in terms of models, and use of science. PISA 2006 competencies were identifying scientific issues, explaining phenomena scientifically and using scientific evidence (PISA, 2007). Table 2.1 showed the important characteristics of three competencies.

Table 2.1 The Competencies for the PISA 2006

Identifying Scientific Issues
• Recognizing issues that are possible to investigate scientifically • Identifying keywords to search for scientific information • Recognizing the key features of a scientific investigation
Explaining Phenomena Scientifically
• Applying knowledge of science in a given situation • Describing or interpreting phenomena scientifically and predicting changes • Identifying appropriate descriptions, explanations, and predictions
Using Scientific Evidence
• Interpreting scientific evidence and making and communicating conclusions • Identifying the assumptions, evidence and reasoning behind conclusions • Reflecting on the societal implications of science and technological developments

Note. From *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006* (p.29), 2006, Paris: OECD. Copyright 2006 by OECD. Reprinted with permission.

Furthermore, the competencies were very important in the definition of science literacy in PISA 2006. In order to be successful in these competencies, students should have both knowledge of natural world and knowledge about science itself for science knowledge and necessary attitudes towards science-related issues.

The students' scores were grouped in six proficiency levels in terms of the achievement in competencies. The students were assigned the highest proficiency level in which they could solve correctly most of the questions. Level two was assigned as baseline proficiency level which means students could use their knowledge in daily life. In Level two, the expected competencies from students were identifying key features of a scientific investigation, remembering single scientific knowledge relating a situation and using the data to comment the ideas about the results of an experiment. In Level one, students have problems in investigation of key features in addition to misunderstandings in using scientific information. In the Preliminary National Report of PISA, it was indicated that 77.9 % of the students are on the Level two or below Level two in science as seen in Table 2.2 (PISA 2006: The Preliminary National Report, 2007).

Table 2.2 The Distribution of Turkish Students' on Proficiency Levels in PISA

	Average Score in science	Below Level 1 (%)	Level 1 (%) 334,5-409.1	Level 2 (%) 409.1-483.8	Level 3 (%) 483.8-558.5	Level 4 (%) 558.5-633.1	Level 5 (%) 633.1-707.8	Level 6 (%) Higher than 707.8
Turkey	424	12,9	33.7	31.3	15.1	6.2	0.9	0.0

Note. From *PISA 2006:The Preliminary National Report* (p. 29), 2007, Ankara: EARGED.

In this report by Ministry of National Education, the competencies of Turkish students were indicated as follows:

- "Turkish students had enough scientific knowledge to describe possible explanations in ordinary situations or deduced the consequences of simple investigations.
- Turkish students could make direct reasoning and interpret the results of a scientific inquiry or technological problem solving (PISA 2006: The Preliminary National Report, p. 29, 2007)."

2.1.1.3 Contexts

PISA 2006 aimed to assess students' preparedness for future life. Therefore, there were many science questions related to daily life situations such as "Health", "Natural Resources", "Environmental quality", "Hazards" and "Frontiers of science and technology". These subjects could be explained in three contexts as personal (the self, family and peer groups), social (community) and global (life across the world). While choosing the contexts, students' interests and lives, the scientific events that adults face with were taken into consideration. Olsen (2004) emphasized that the content of items were extracted from daily life materials such as popular science magazines, newspapers, and advertisements.

In Table 2.3, the contexts in PISA were given.

Table 2.3 The Contexts for the PISA 2006

	Personal (Self, family and peer groups	Social (The community)	Global (Life across the world)
Health	Maintenance of health, accidents, nutrition	Control of disease, social transmission, food choices, community health	Epidemics, spread of infectious diseases
Natural Resources	Personal consumption of materials and energy	Maintenance of human populations, quality of life, security, production and distribution of food, energy supply	Renewable and nonrenewable, natural systems, population growth, sustainable use of species
Environment	Environmentally friendly behavior, use and disposal of materials	Population distribution, disposal of waste, environmental impact, local weather	Biodiversity, ecological sustainability, control of pollution, production and loss of soil
Hazard	Natural and human induced, decisions about housing	Rapid changes (earthquakes, severe weather), slow and progressive changes (coastal erosion, sedimentation), risk assessment	Climate change, impact of modern warfare
"Frontiers of science and technology"	Interest in science's explanations of natural phenomena, science based hobbies, sport and leisure, music and personal technology	New materials, devices and processes, genetic modification, transport	Extinction of species, exploration of space, origin and structure of the universe

Note. From *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006* (p.27), 2006, Paris: OECD. Copyright 2006 by OECD. Reprinted with permission.

2.1.1.4 Attitudes

Besides assessing the scientific and technical knowledge, developing interest in science and supporting for scientific enquiry were very important. Moreover, attitudes towards science were very effective in students' decisions of developing the science knowledge or using scientific concepts in their daily life. The attitudes for PISA were built upon the reviews of attitudinal research. The data gathered on attitudes were mainly in four areas: support for scientific enquiry, self-belief as science learners, interest in science and responsibility towards resources and environments as seen in Table 2.4.

Table 2.4 The Assessment Areas of Attitudes for the PISA 2006

Support for Scientific Enquiry
• Acknowledge the importance of considering different scientific perspectives and arguments • Support the use of factual information and rational explanations • Express the need for logical and careful processes in drawing conclusions
Self-belief as Science Learners
• Handle scientific tasks effectively • Overcome difficulties to solve scientific problems • Demonstrate strong scientific abilities

Table 2.4 (continued) The Assessment Areas of Attitudes for the PISA 2006

Interest in Science
• Indicate curiosity in science and science-related issues and endeavors
• Demonstrate willingness to acquire additional scientific knowledge and skills, using a variety of resources and methods
• Demonstrate willingness to seek information and have an ongoing interest in science, including consideration of science-related careers
Responsibility towards Resources and Environments
• Show a sense of personal responsibility for maintaining a sustainable environment
• Demonstrate awareness of the environmental consequences of individual actions
• Demonstrate willingness to take the action to maintain natural resources

Note. From *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006* (p. 37), 2006, Paris: OECD. Copyright 2006 by OECD. Reprinted with permission.

In PISA, these above mentioned four areas were given priority due to giving a general idea about students' appreciation of science and reflecting students' levels of engagement of all students in science. One of the attitudes including gathering evidence, reasoning rationally, responding critically and communicating conclusions when students face with science topics in their life was "Support for Scientific Enquiry". Another attitude was "Self Belief" reflecting students' own abilities as a part of achievement in science. The reason of selecting "Interest in Science" was the results of studies showing that this attitude was a predictor of later science learning such as students' consideration of science related careers. The last attitude in PISA 2006 was about a global problem namely "Responsibility towards Resources and Environments". This gave information about students' level of concern in environmental issues.

2.2 Science Assessment Framework in TIMSS

2.2.1 The Development Process for the TIMSS 2007 Assessment Framework

In TIMSS 2007, the TIMSS & PIRLS International Study Center put forth effort to update the framework for the TIMSS 2003. U.S. National Science Foundation was also in this process in order to enable appropriate frameworks for the countries participating TIMSS. In TIMSS 2007, there were specific objectives in science and mathematics for assessing students for the first time. In developing the framework of TIMSS, an international panel of mathematics and science was held and experts in test construction guided general form of framework. Also the representatives from national centers contributed in developing this framework. The participant countries completed the questionnaires about the curricula in order to have consistent goals of mathematics and science in a significant numbers of countries. The questionnaire was giving feedback about the topics if they were suitable in terms of content and behaviors for the fourth and eighth grades in their countries. Then, the framework was revised by Science and Mathematics Item Review Committee (SMIRC) based on the feedback provided by participating countries. The final review was done by National Research Coordinator (NRC) before publication.

Science Assessment Framework organized in three dimensions namely content domain, cognitive domain or thinking processes and contextual domain in TIMSS 2007.

2.2.2 Content Domain

The content domain was different for fourth and eighth grades. In eighth grade, four content domains namely biology, chemistry, physics, and earth science and in fourth grade, three content domains namely life science, physical science and earth science were included in structure of TIMMS 2007. While life science was given more importance in fourth grade, on the other hand, physics and chemistry were more important than the biology in eighth grade.

2.2.3 Cognitive Domain

In TIMSS, there were three cognitive domains in both fourth and eighth classes. In the Table 2.5, these three domains with the expected behaviors were given.

Table 2.5 Cognitive Domain for the TIMSS 2007

Cognitive Domain	Expected Behaviors
Knowing	Recalling-Recognizing Defining Describing Illustrating with examples Using tools and procedures

Table 2.5 (continued) Cognitive Domain for the TIMSS 2007

Cognitive Domain	Expected Behaviors
Applying	Comparing, contrasting and classifying
	Using models
	Relating
	Interpreting information
	Finding solutions
	Explaining
Reasoning	Analyzing and solving problems
	Integrating and synthesizing
	Hypothesizing and predicting
	Designing and planning
	Drawing conclusions
	Generalizing
	Evaluating
	Justifying

Note. From *TIMSS 2007 Assessment Framework* (pp. 69-75), by Mullis I. V. S. et al., 2005, Boston College: TIMSS & PIRLS International Study Center, Lynch School of Education. Copyright 2005 by the International Association for the Evaluation of Educational Achievement (IEA). Reprinted with permission

The first domain was “Knowing” which included covering facts, procedures and concepts. Furthermore, “Knowing” was required for engaging successfully to more complex cognitive activities. “Applying”, the second domain, included covering the ability to apply the knowledge and conceptual understanding in problematic situations. The third domain was “Reasoning” which referred that students’ ability to find solutions in unfamiliar situations, complex contexts and multi-step problems. This domain required more complex behaviors than the other domains. The percentage of cognitive domains differed in grades because of the maturity, different cognitive ability, experience, the understanding levels of students and instruction.

In TIMSS, in order to make logical comparisons between countries and students, the performance was assessed via the science achievement scale which was based on the achievement on content and cognitive domains. There were four scales in TIMSS 2007 as The Advanced International Benchmark, the High International Benchmark, the Intermediate International Benchmark and the Low International Benchmark as seen in Table 2.6.

Table 2.6 The Distribution of Turkish Students on TIMSS 2007 Benchmarks

TIMSS	Advanced B. (%) 625	High B. (%) 550	Intermediate B. (%) 475	Low B. (%) 400
Turkey	3	16	40	71

Note. From *TIMSS 2007 International Science Report* (p. 69), by Martin M.O. et al., 2008, Boston College: TIMSS & PIRLS International Study Center, Lynch School of Education. Copyright 2008 by the International Association for the Evaluation of Educational Achievement (IEA). Reprinted with permission.

Since most of the Turkish students were at Low Benchmark, the characteristics of this benchmark were explained:

- "Students could recognize some basic facts from the life and physical sciences.

- Students had some knowledge of the human body, and demonstrate some familiarity with everyday physical phenomena.
- Students can interpret pictorial diagrams and apply knowledge of simple physical concepts to practical situations (TIMSS 2007 International Science Report, p. 108, 2008)."

2.2.4 Contextual Domain

The third domain was contextual domain which was investigating the types of factors and contexts affecting students' learning.

2.3 Studies about the Comparison of PISA & TIMSS Frameworks

PISA and TIMSS were both large scale international studies. There were several countries participating more than one international study and students have different scores on the assessments. Therefore, it was useful to compare the frameworks of international studies to explain the difference. While there were many studies (Bybee, 2007; Geske & Kangro, 2002; Ginsburg et al., 2005; Harlen, 2001; Hutchison & Schagen, 2007; Lie, 2005) comparing PISA and TIMSS at the level of broad aims, there wasn't enough studies comparing at the item level in science (Dossey, 2006; Nohara, 2001; Ruddock et al; 2006; Wu, 2008). Moreover, some of the studies claimed that there could not be a direct comparison between TIMSS and PISA (Adams, 2003; Harding, 2003; Olsen, 2004).

Bybee (2007) analyzed PISA and TIMSS 2003 in a general way including sampling, organization, content, assessment cycle and aim. Bybee reported one of the benefits of international studies as giving clues about how to be a high achiever country. Also, international studies have enabled the chance of comparing the countries school systems with high achiever countries. Bybee pointed out that USA students were more successful in curriculum based studies such as TIMSS or NAEP but not in PISA which was context-based assessment. Therefore, the author concluded that there should be reforms on school system of USA in order to cope with economic competitors (Bybee, 2007).

Another study comparing PISA and TIMSS in a general way was the study of Ginsburg et al. (2005). Ginsburg et al. explained that PISA aimed to assess to apply scientific ideas to real-world problems whereas TIMSS aimed to assess the knowledge of the curriculum taught in schools. In the study, another difference between PISA and TIMSS was the percentages of multiple choice or constructed response questions. (Ginsburg et al., 2005).

Hutchison and Schagen (2007) also studied on comparison of TIMSS and PISA in terms of aims, sampling and items. While TIMSS' aim was to investigate what's happening in the classroom, PISA aimed to find what will happen in adult life of students regarding required literacy. PISA items were investigating if students could relate among contexts, while TIMSS items were evaluating to have scientific knowledge. Another difference was in analyzing data in TIMSS and PISA. While in PISA, it has been used 1-parameter Rasch-model, TIMSS has been using 2-parameters Item Response Theory Model. In conclusion, Hutchison and Schagen (2007) added that there should be more analysis on PISA and TIMSS about their apparent discrepancies because in the studies there were clues of how a national system would have a high quality education (Hutchison & Schagen, 2007).

Harlen (2001) assessed the framework of scientific literacy in PISA and compared PISA 2000 with TIMSS-R. He declared one of the most significant differences between PISA and TIMSS in the following way. PISA was seeking what to be learned for future life and TIMSS was investigating if the common core goals of curriculum were learned. According to Harlen (2001), the most significant difference between PISA and TIMSS was the design aims. While TIMSS was assessing the progress of students in different ages for countries, PISA was concerned with the outcomes of the compulsory education and if students would use processes and skills in future life of citizens rather than just memorizing facts (Harlen, 2001).

Olsen (2004) presented the study of the assessment about scientific literacy of PISA at NARST. The author declared the scientific literacy in PISA 2006 was centered at scientific/cognitive competencies rather than very detailed expected behaviors. The reason for that was explained as follows. Since people were more familiar with scientific processes, PISA used scientific process skills in the scientific literacy. Furthermore, in the study, TIMSS and PISA were compared. Scientific Literacy in PISA was different from the science concept in TIMSS in terms of assessing what adolescents would need in the present and future life. On the other hand, TIMSS was assessing the country achievement what learnt at school. In other words, Olsen (2004) declared that TIMSS was based on the curriculum analysis. TIMSS measured how successful the school system in applying the curriculum policy. PISA evaluated how successful the curriculum and school system enabled students' basic competencies in order to be good citizens in future.

In another study of Olsen (2005), the differences between TIMSS and PISA were analyzed in a general way. TIMSS 8^{th} grade students were in general two years younger than PISA students. This stemmed from the case that TIMSS was grade based, while PISA was age based study. Therefore, the grades of students in PISA were different in countries. Another difference was that sampling procedures of TIMSS and PISA was different. TIMSS sampled classes within schools but PISA sampled students within schools. The questionnaires in TIMSS and PISA had different aims too. In PISA, questionnaires were trying to collect information about students' social background. In TIMSS, the aim was to collect information about the teaching of the subject matter (Olsen, 2005).

There are several studies comparing PISA and TIMSS but Lie (2005) emphasized on that the aim of the comparison was not to make the copies of successful countries' education system. The author considered the comparison of PISA and TIMSS would provide countries to improve their educational system by taking into account their own culture, historical and pedagogical traditions. He explained the most distinct difference between PISA and TIMSS framework in the following way. PISA measured how well students were ready for the future life, in other words the preparedness for life. On the other hand, TIMSS measured the achieved curriculum. While TIMSS was curriculum driven, PISA was utility-relevance driven. PISA tried to investigate which information will help the 15 year students in their future life (Lie, 2005).

Geske and Kangro (2002) looked the reasons of achievement difference between PISA 2000 and TIMSS 1999 in Latvia. While Latvia had one of the lowest scores in PISA, in TIMSS Latvia had higher scores than nearly half of the countries. This result was explained as following. The frameworks of two assessments differed which the common curriculum of school science was used in TIMSS by consulting to all participant countries in TIMSS; in PISA, scientific literacy was tested in other words, students' ability to conclude a decision using scientific skills in daily life. Another difference was about the types of questions. In TIMSS most of the questions were multiple choice questions or requiring short answers but in PISA, there were a long page questions requiring reading skills (Geske & Kangro, 2002).

Although most of the researchers compared PISA and TIMSS in a general way, there are studies analyzing these international studies in item level. One of these studies was the study of Nohara (2001). This study was about the comparison of PISA 2000, TIMSS-R and National Assessment of Educational Progress (NAEP) frameworks for eighth grades. Nohara's study was based on the work of expert panels in science and mathematics by reviewing individual items in the assessments. Panel experts emphasized on the following three questions while comparing three assessments. These questions were if the assessments cover the same topics, if the type of asking questions were the same and if the assessments used the similar type of thinking skills were questioned. It was found that in TIMSS-R, the number of Physical Science items was more than NAEP and the number of items among contents in PISA was more equally distributed than TIMSS-R but less than NAEP. Another finding of the study was the cognitive processes were more emphasized in PISA rather than TIMSS-R or NAEP. The most common type in the assessments was multiple choice questions. TIMSS-R had more free response and extended free response questions than PISA 2000. In the study, the percentage of real world questions among three assessments was also questioned. As expected, PISA 2000 with 66 % had the most real life items. NAEP with 23 % was the second assessment while TIMSS-R had the 16 % of science items relevant to practical situations. It has found that mathematical skills in science items were in most PISA with 20 %. Nohara (2001) defined multi-step reasoning as the ability to construct information in order to solve the question and it was found that 77 % of science items in PISA, 44 % of science items in NAEP and 31 % of science items in TIMSS included multi-step reasoning. Nohara (2001) also declared that comparison of three assessments showed that there were differences in the purposes and philosophical underpinnings of the studies. While TIMSS-R and NAEP had a purpose of finding the proficiency level of students in science and math content, PISA aimed the ability of students to use the scientific and mathematics skills in non-school environment. Another finding from the panel experts was the difference of the structure, content and nomenclature in the assessments (Nohara, 2001).

Ruddock et al. (2006) also compared PISA and TIMSS in the item level. In the study, the effect of differential item familiarity on the performances of English students in TIMSS and PISA was investigated. Experts did rating study in terms of content or skills, the contexts, item format and overall appropriateness for the PISA 2000 and 2003 and TIMSS 1999 and 2003 eighth grade science items. The method was answering the familiarity for concept, the context and the format of each question. In the concept and the context, the familiarity rate was over 50 %, while in item format it was over 40 % both in TIMSS and PISA. Another finding was that the grammatical complexity in science items was found to be higher in PISA than in TIMSS (Ruddock et al., 2006).

In the study of Dossey et al. (2006) the major focus was the comparison of problem solving in TIMSS 2003 and PISA 2003. Dossey et al. (2006) compare these two international studies in the item level. The items included problem solving skills with the percentage of 26 % in TIMSS and 49 % in PISA. The analyses showed that there was not so much difference in allocation of science items in content and cognitive domains of assessments in the problem solving items. Moreover, it was found that the percentage of science items demanding “critically evaluate information” was higher in PISA but the science items that “required science knowledge” items were higher in TIMSS. The most striking difference between TIMSS and PISA was the format of the items. The problem solving items in TIMSS was mostly multiple choice questions, on the other hand, PISA had close short constructed response format for problem solving items (Dossey et al., 2006).

Another study of comparison in item level in science was by Wu (2008). The author studied the achievement levels of 22 participating countries in TIMSS and PISA 2003 in science and mathematics. She used country mean scores and regression techniques in comparing countries in addition to two variables (years of schooling in PISA and content balance) to explain the observed differences in countries' performance. The correlation between PISA and TIMSS science country mean scores was found as 0,95 and the R^2, the proportion of variance was found as 0,90 which was showing 90 % of TIMSS Science scores was explained by PISA scores. When the TIMSS age and years of schooling predictors added, R^2 was increased to 92 %. It was assumed that such a high proportion of variance stemmed from the small variations of content area scores within each country in science. In other words, countries had similar strengths and weaknesses in science content area. Another finding of the study was that the number of schooling years and the achievement was proportional to each other. Moreover, it was found that the age of students was less effective than the years of schooling (Wu, 2008).

Harding (2003) was one of the researchers saying there could not be direct comparison between PISA and TIMSS. The author explained its reason as the different target population in PISA and TIMSS 2003 but Harding also made a comparison between PISA and TIMSS in terms of purposes and frameworks. In TIMSS, the assessment was based on the framework from curricula; on the other hand, PISA emphasized which skills and competencies 15 year students can apply in the real world context. By focusing on the application of data and mastery of processes, PISA also underlined learning at outside of the school (Harding, 2003).

Another researcher asserting that direct comparison wouldn't be done between PISA and TIMSS was Olsen (2004). The author stated that PISA and TIMSS differed in their assessing aims. This was the reason why there could not be any direct comparison between PISA and TIMSS. Another reason for that explained in the study of Olsen (2005) as the difference of measured scientific concepts in two studies. Then, he made a comparison between PISA and TIMSS. The first difference was the test domain between PISA and TIMSS. In PISA, the curriculum was not explicitly covered in the framework although in TIMSS, framework aimed to represent national curricula. Another difference was the population and the way of sampling in the studies. While in TIMSS there were two class based population, in PISA there was one age based population. The design and instruments also differs in PISA and TIMSS. TIMSS was 4 years cycle test with 50/50 test time for the math and the science but PISA was 3 years cycle test with the rotation of main subject in each cycle as reading, math and science. The last difference defined by Olsen (2004) was the organization and the participants. PISA and TIMSS were organized by different organizations with different countries participating. While TIMSS was organized by IEA, PISA was organized by OECD (Olsen, 2004).

Adams (2003) indicated that the meaning of the PISA and TIMSS scale both was 500 and the standard deviation of the scores was 100. Although it seemed that two international studies were comparable, the participating countries and the target populations from each participant countries were different. Therefore, PISA and TIMSS were not statistically linked and comparable (Adams, 2003).

There were not so much studies emphasizing the analysis of frameworks of PISA and TIMSS deeply. The studies mentioning about competencies in PISA and cognitive skills in TIMSS in detail were given as follows.

Ciascai (2009) compared Romanian curriculum with TIMSS 2007 via Bloom's Taxonomy of cognitive domain. In the analysis, the author used the number and type of action word used in formulating the cognitive skills. In the analysis, it was found that there were more questions of comprehension skill, then application and synthesis questions. Then, the author analyzed the Romanian curriculum and compared with TIMSS 2007. There were a low percentage of common cognitive skills in TIMSS 2007 and Romanian curriculum. According to the author, this was the reason why Romanian students presented poor performance in TIMSS 2007 (Ciascai, 2009).

Kjærnsli and Lie (2004) studied the difference between girls and boys on competencies as demonstrating understanding of scientific concepts vs. process skills in scientific reasoning for Nordic countries in PISA 2000. These process skills were recognizing scientifically investigable questions, identifying evidence needed in a scientific investigation, drawing or evaluating conclusions and communicating valid conclusions. Process items were found more difficult according to the results of countries. It was found that there was no clear pattern of gender differences in science in PISA 2000 (Kjærnsli & Lie, 2004).

In another study, Kjærnsli (2003) explored the science performance of students in scientific literacy of Nordic countries including Norway, United States, France, United Kingdom, Denmark, Finland, Sweeden, Germany, Russia, Iceland and Hungary. The author defined science performance under two categories in PISA 2000 like in the study of Kjærnsli and Lie (2004). One of them was Scientific Process Skills which were intellectual processes used in solving a science question and the other was content dimension which was scientific knowledge. In PISA 2000, Kjærnsli (2003) found that Nordic countries had higher scores in the items concerning content dimension than the items concerning science process skills. He claimed the reason as the difficulty of science process skills (Kjærnsli, 2003).

Although there were studies claiming that PISA and TIMSS could not be compared, there were the others comparing PISA and TIMSS. Another important point was the similarities between scientific process skills with competencies in PISA.

CHAPTER 3

METHODS

This chapter contained the methods of the study including PISA and TIMSS population and sample, instruments, reliability and validity of the studies, data collection and data analysis.

3.1. Population and Sampling

Since PISA was an international study, the comparability of the results over countries gained more importance. The schooling age and the structure of education system differed in countries. Therefore, the population should be well-defined with reference to a target age rather than school grades for a valid international comparison of educational performance. The target population was the students between 15 years 3 months and 16 years 2 months having completed at least 6 years of formal schooling regardless of the type of institution for PISA.

In sampling, explicit stratification by region for total of seven explicit strata was done in Turkey for both TIMSS and PISA. Sampling process in PISA and TIMSS was very similar to each other. In both assessments, two stage stratified sample design was used. While in the first stage, in both assessments the schools were selected, the second stage differed. In PISA, students were selected while in TIMSS, classrooms were selected from the sampled schools.

A two stage stratified sample design was used in PISA 2006. In the first stage sampling, only the schools having 15 years old students were chosen with probabilities proportional to a measure of size. In other words, systematic probability proportional to size (or PPS) sampling was used which measure of size was estimated number of eligible 15 year old students enrolled. In the second stage, students from sampled schools were the sampling units. After a list of 15 year students from the sampled schools, the target cluster (the TCS) size was determined as 35. Then from the list more students than TCS were selected. Students were the sampling units in the third stage of sampling. Each country's National Project Manager should use KeyQuest, the PISA Sampling Software, to select student samples. An international consortium, referred to as the PISA Consortium, was responsible for the design and implementation of PISA 2006, led by the Australian Council for Educational Research (ACER) and by the other partners in the consortium including the Netherlands National Institute for Educational Measurement (CITO), the National Institute for Educational Policy Research in Japan (NIER) and WESTAT in the United States. In PISA 2006, the major focus was science (PISA Technical Report, 2009).

TIMSS was an international study assessing students at fourth and eighth grades. It would have some difficulties to carry out an international grade-based test because of the different entry ages to primary schools. Therefore, TIMSS specified the target population to prevent testing of very young children. According to TIMSS guideline, the minimum age of fourth grade should be 9.5 years old and the maximum age of eight grades should be 13.5 years old. If there were countries starting primary school to early, they would be assessed at the next grade. The aim of grade based study was explained as to assess students' achievement after the same amount of schooling. A two-stage probability proportional to size sampling design (or PPS) was used in TIMSS 2007. While sampling of schools was done in the first stage, sampling of intact classrooms was selected with equal probability of selection using systematic random sampling from the selected schools. In general, participating countries selected 150 schools and one or two intact classrooms from these schools. In the representative sample, there were at least 4500 students in each country. Random selection of schools was carried out the Netherlands, Scotland and England but also Canada controlled the sampling.

The Science and Mathematics Item Review Committee (SMIRC) was an international committee who was responsible for reviewing and revising items, checking for mathematical and scientific accuracy and making certain of congruency between defined specifications and items. National Research Coordinators (NRCs) with the support and training from TIMSS & PIRLS International Study Center wrote items in large measure (TIMSS Technical Report, 2008).

Since the studies for PISA 2006 and TIMSS 2007 occurred at the same time, overlap control procedure for requesting eight countries (Australia, Bulgaria, England, Hong Kong-China, Hungary, Scotland, Tunisia, and the USA) was used. This application stemmed from the potential for the increased burden. Since TIMMS samples were selected before PISA, the TIMSS and PIRLS International Study Center gave the PISA Consortium sampled schools for the requesting eight countries.

3.2 Instruments

3.2.1 PISA

In PISA 2006, there were achievement test, student, school and parent context questionnaires. These instruments were prepared in English and translated into 44 languages in 87 national versions.

3.2.1.1 Cognitive Tests:

In PISA 2006, three domains were examined namely science, mathematics and literature. In this study, scientific literacy discussed. There were 37 science units including 108 cognitive items which 22 items were from PISA 2003 and 31 embedded attitudinal items. There were 13 booklets and each booklet contained 4 clusters of questions. There were 7 science, 4 mathematics and 2 reading clusters. Since science was the major domain, in every booklet there was science cluster. Booklets were randomly selected for students and students were given 120 minutes for PISA 2006. The distribution of science items in the scientific competency and science content areas were shown in the Table 3.1 and Table 3.2.

Table 3.1 Science Main Study Items (Item Format by Competency)

Competencies	Item Format		Total
	Multiple-choice	Constructed response	
Identifying scientific issues	19	5	24 (22%)
Explaining scientific phenomena	33	20	53 (49%)
Using scientific evidence	15	16	31 (29%)
Total	67 (62 %)	41 (38 %)	108

Note. From *PISA 2006 Technical Report* (p. 43), 2009, Paris: OECD. Note. From *PISA 2006 Technical Report* (p. 43), 2009, Paris: OECD. Copyright 2009 by the Organisation for Economic Co-operation and Development (OECD). Reprinted with permission.

Table 3.2 Science Main Study Items (Item Format by Science Content Area)

Science Content Area	Item Format		Total
	Multiple-choice	Constructed response	
Knowledge of Science Physical Systems	11	6	17
Knowledge of Science Living Systems	16	9	25
Knowledge of Science Earth and Space Systems	7	5	12
Knowledge of Science Physical Systems	11	6	17

Table 3.2 (continued) Science Main Study Items (Item Format by Science Content Area)

Science Content Area	Item Format		Total
	Multiple-choice	Constructed response	
Knowledge of Science Technology Systems	5	3	8
Knowledge about Science Scientific Enquiry	19	6	25
Knowledge about Science Scientific Explanations	9	12	21
Total	67	41	108

Note. From *PISA 2006 Technical Report* (p. 43), 2009, Paris: OECD. Copyright 2009 by the Organisation for Economic Co-operation and Development (OECD). Reprinted with permission.

After two hours cognitive test, students were given a 30-minute questionnaire for assessing the attitudes towards science and science school teaching and also another questionnaire for providing information about computer access and frequency of usage of computers were given in Turkey. Another questionnaire was about parents' investment on their children and views on science related issues which was completed by parents.

3.2.1.2 Validity and Reliability

An international consortium, referred to as the PISA Consortium, was responsible for the design and implementation of PISA 2006, led by the Australian

Council for Educational Research (ACER) and by the other partners in the consortium including the Netherlands National Institute for Educational Measurement (CITO), the National Institute for Educational Policy Research in Japan (NIER) and WESTAT in the United States. After items were reviewed by Expert Groups, the guide to coding of responses for each question was prepared. Also overview of development process of items and the importance of framework fit were placed in the guide. Through development process, coding guide and units in standard format were prepared in the local language. Then, item paneling including analyzing the items from the point of view of a student and a coder, cognitive interviews with individual students and group of students and finally local pilot study were done. With the feedbacks from each process, items were reviewed. Then, these items were reviewed by another development team which was not one of the teams in the first phase and items were again tested in different pilot schools in Australian schools. The participant countries also rated the items in terms of item difficulty, cultural appropriateness and coverage of items in the curriculum. Thus, validity of the items was ensured.

In PISA, it was underlined that the test had to be done in the language that students have instructed. Therefore, the items were translated into 44 languages. In the translation process, after translating into the language that would be tested, the items were again translated into English to find out discrepancies. Then, a reconciler checked if there were any differences with the original source. The consortium verifier also checked the national source against the English and French source.

Administration process was very well-defined in PISA. A National Project Manager (NPMs) in each participating country was responsible of implementing PISA. Besides the NPMs, the roles of several assistants, school level staff namely school co-ordinators (SCs) and trained test administrators were defined. For high reliability, National Project Managers were given the following guides; A National Project Manager's Manual, A School Sampling Preparation Manual and A Data Management Manual and also A School Co-ordinator's Manual for School Co-ordinators and Test Administrator's Manual for test administrators were given.

The reliability analysis was done for five scales including science, mathematics, reading, the attitude scales, interest and support and the values were given in Table 3.3.

Table 3.3 Reliability Values in PISA 2006

Domain	Reliability (Cronbach's Alpha)
Science	0.91
Explaining phenomena scientifically	0.90
Identifying scientific issues	0.90
Using scientific evidence	0.92

3.2.2 TIMSS

In TIMSS 2007, besides achievement tests of science and mathematics, there were four background questionnaires namely curriculum, school, teacher and student questionnaires.

The curriculum questionnaires were designed to provide content and implementation of countries' curriculum for science and mathematics at fourth and eighth grade. Another purpose of designing was to check if the topics were parallel to countries' intended curriculum. In the questionnaire, it has also emphasized the policies of parental involvement and assigning homework.

Another background questionnaire namely school questionnaire was completed by administration in thirty minutes like in PISA and its aim was to collect information about the factors influencing student achievement in mathematics and science.

Furthermore, teacher questionnaire was designed to provide information about teaching and learning environment. From each participating schools, just one class' science teacher(s) would complete the questionnaire in 45 minutes.

Every student participating TIMSS should complete the 30-minute student questionnaire. It gathered information about students' background, attitudes towards science and mathematics, resources for learning and learning experiences.

3.2.2.1 Achievement Test:

At the eighth grade, there were 429 items including 214 science items and 215 mathematics items in TIMSS 2007. There were 23 science items from TIMSS 1999 and 84 science items from TIMSS 2003, while 133 science items were new in TIMSS 2007. There were 14 booklets and in each booklet, there were 2 blocks of science items and 2 blocks of mathematic items. In TIMSS 2007, for the eighth grade 22.5 minutes were given for each block so as total 90 minutes was given for each booklet to students. The number of questions in TIMSS 2007 was given in terms of item format in Table 3.4.

Table 3.4 Science Main Study Items by Item Format

Content Domain	Item Format		Total
	Multiple Choice	Constructed Response	
Biology	36	40	76
Chemistry	21	21	42
Physics	31	24	55
Earth Science	19	22	41
Total Science Items	107	107	214

Note. From *TIMSS 2007Technical Report* (p. 40), by Olson J. F. et al., 2008, Boston College: TIMSS & PIRLS International Study Center, Lynch School of Education. Copyright 2008 by the International Association for the Evaluation of Educational Achievement (IEA). Reprinted with permission.

The distribution of score points including both content and cognitive skills in TIMSS 2007 were given in Table 3.5.

Table 3.5 The Distribution of Score Points in Science Main Study Items by Content and Cognitive Domains in TIMSS

Content Domain	Cognitive Domain			Total
	Knowing	Applying	Reasoning	
Biology	35	31	23	89 (37 %)
Chemistry	16	18	12	46 (19 %)
Physics	14	32	13	59 (25 %)
Earth Science	24	16	4	44 (19 %)
Total score points	89 (37 %)	97 (41 %)	52 (22 %)	238

Note. From *TIMSS 2007Technical Report* (p. 40), by Olson J. F. et al., 2008, Boston College: TIMSS & PIRLS International Study Center, Lynch School of Education. Copyright 2008 by the International Association for the Evaluation of Educational Achievement (IEA). Reprinted with permission.

3.2.2.2 Validity and Reliability:

TIMSS 2007 Assessment Framework was reviewed and revised by the international committee called the Science and Mathematics Item Review Committee (SMIRC) in which mathematics and science experts from different nations and cultures. National Research Coordinators, the TIMSS & PIRLS International Study center staff, the Mathematics and Science coordinators and the consultants from Educational Testing Service analyzed the items in TIMSS 2003 and made some revisions for the items in TIMSS 2007. Then, NRC's were met to update the proposed items and recommended on the content and cognitive domains. NRCs were asked to give the specific topics in science and mathematics by a questionnaire given by TIMSS &PIRLS International Study Center. In the second NRC meeting, new multiple choice and constructed response items were written in four subgroups as mathematics and science in the fourth grade and eighth grade in the light of instructions of TIMSS &PIRLS International Study Center. Items then reviewed in two groups as science and mathematics in terms of content accuracy, framework fit and grade appropriateness by SMIRC. In the third meeting of NRCs, the items were reviewed and controlled in terms of content accuracy, framework fit and grade appropriateness too. The, newly developed items were field tested in 45 participant country and in 25 schools for each country in the eighth grade. The results from the field test determined the items used in TIMSS 2007. The TIMSS & PIRLS International Control Center, the mathematics and science coordinators, then SMIRC reviewed the selected items in terms of range of difficulties and coverage of mathematics and science items.

Another important point in validity was bout the translation of items into the language of participant countries' language. The items were translated into 39 languages by experienced translators. Translators had to be qualified in English language, the target language, also experienced in subject matter in addition to countries' cultural context. A translation reviewer controlled the translated items.

Following the control of translation of the test instrument, National Research Coordinators controlled and analyzed suggestions. After that a verification process continued with internal verification of the translations at the national centers, independent verification by an international translation company and a check by International Quality Control Monitors to check if the verifiers' suggestions has been taken into consideration or not. Finally, the TIMSS &PIRLS International Study Center analyzed the translated instruments from all participating countries.

There were adequate questions to ensure reliable measurement and also there was a guidebook explaining the administration process and also scoring procedures. Another caution was the training which helped in providing that the results were not affected from extraneous factors.

The role of the National Research Coordinators was so valuable because the NRCs was responsible of implementing TIMSS in their countries. TIMSS & PIRLS International Study Center, IEA Secretariat, IEA Data Processing and Research Center (DPC), and Statistics Canada developed an internationally standardized survey operations procedures. Every step in the administration process was explained in that guideline.

For ensuring reliability, the same the scoring guideline from TIMSS 2003 for constructed response items were used in TIMSS 2007. The field test of the constructed response items for clarifying the scoring guideline was done in English speaking countries. Besides the data, feedback from the trainers also was provided such as experiences during scaling.

The science reliability (Cronbach's alpha) was found as 0.84 for eighth grade in overall and the reliability was 0.85 in Turkey for the eighth grade in TIMSS 2007.

In TIMSS, it was important to reach a decision such the students were really good at the domains which they had high achievement. In other words, it has given so much importance to the validity of the tests.

Besides the validity of the instrument, the comparable validity was also important because the differences in the achievement of the countries should be valid. Therefore, preparing of the TIMSS framework, developing the instruments and operational procedures and translation verification process became more important.

3.3 Data Collection

In the present study, data collection did not take place because PISA and TIMSS data was used. At first, PISA data collection was explained.

In Turkey, the PISA 2006 was done in May 2006. Turkey participated PISA 2006 with the 160 schools and 4942 (2290 females, 2652 males) students. In the Table 3.6 and 3.7, the number of students from different school types participating PISA and the geographical regions of students with the gender information were given.

Table 3.6 The Number of Students by School Types in PISA 2006

School Types	Number of Students	Percentage of Students (%)
Elementary Schools	116	2.3
Public High Schools	2266	45.9
Anatolian High Schools	549	11.1
Super Lycee*	9	0.2
Science High Schools	35	0.7
Vocational Schools	1510	30.6
Anatolian Vocational Schools	179	3.6
Multi-programme High School	278	5.6
TOTAL	4942	100.0

Note. From *PISA 2006 the Preliminary National Report* (p. 12), 2007, Ankara: EARGED.

*In Super Lycee, the lessons were given in English.

Table 3.7 The Number of Students by Geographical Regions in PISA 2006

Geographical Regions	Number of Girls	Number of Boys	Total Number of Students
Marmara	720	718	1438
Agean	298	324	622
Mediterrian	360	348	708
Central Anatolia	396	469	865
Black Sea	220	376	596
Eastern Anatolia	185	169	354
Southeast Anatolia	111	248	359
TOTAL	2290	2652	4942

Note. From *PISA 2006 the Preliminary National Report* (p. 13), 2007, Ankara: EARGED.

A school coordinator was chosen for listing 15 year-old students. Then, the responsible person in implementing PISA 2006 was chosen by PISA Consortium. That person was responsible of distributing the booklets correctly, collecting the booklets and sending them to PISA Consortium. The responsible person read the instructions from a paper to provide unity in each participant country.

In Turkey, the TIMSS 2007 was done in April 2007. Turkey participated TIMSS 2007 with the 146 schools and 4498 students (2384 male, 2114 female) at eight grades as given in Table 3.8.

Table 3.8 The Number of Schools by Geographical Regions in TIMSS 2007

Geographical Regions	Number of Schools
Marmara	39
Agean	18
Mediterrian	18
Central Anatolia	26
Black Sea	16
Eastern Anatolia	12
Southeast Anatolia	17
Total	146

Note. From *TIMSS 2007Technical Report* (p. 423), by Olson J. F. et al., 2008, Boston College: TIMSS & PIRLS International Study Center, Lynch School of Education. Copyright 2008 by the International Association for the Evaluation of Educational Achievement (IEA). Reprinted with permission.

3.4 Data Analysis:

In this study, content analysis of PISA and TIMSS were done in order to find out if the frameworks were parallel. Common items were investigated in PISA and TIMSS. Furthermore, the percentage of correct responses of Turkish students was compared in PISA and TIMSS for the common items to reveal similarities of subscales in two assessments.

3.5 Limitations of the Study

Since all items were not released in PISA and TIMSS, only the released items were used. While selecting common items from PISA and TIMSS to show the similarity, the number of items was limited. With more items, there would more comprehensive results.

CHAPTER 4

RESULTS

This chapter was devoted to the presentation of the present study. Two main sections namely the framework comparison and comparison of competencies by item analyses. In the framework comparison, TIMSS and PISA were compared through content analysis. In the second section, competencies in PISA and cognitive skills in TIMSS were compared with the items representing subscales of PISA and TIMSS.

4.1 Content Analysis of PISA and TIMSS

Several studies have stated that PISA 2006 and TIMSS 2007 differs in survey cycle, age of students assessed, test time, item types, philosophical underpinnings and aims (Bybee, 2007; Ginsburg, 2005; Harding, 2003; Harlen, 2001; Hutchison & Schagen, 2007; Lie, 2005) Another difference was subscales of science in PISA 2006 and TIMSS 2007. In this study, the content analysis was done in comparing PISA 2006 and TIMSS 2007. The comparison was done mainly in two parts as content and the expected behaviors while solving a science problem.

4.1.1 Knowledge in PISA vs. Content Domain in TIMSS

In PISA 2006, knowledge included knowledge of science and knowledge about science.

Knowledge of science was defined as to understand the basic scientific concepts, principles and theories. Since one of the aims of PISA is to assess if the students can use their knowledge in daily life, the topics in PISA were selected according to relevance to real-life situations, compatibility of 15 year-old students' development and importance of scientific knowledge. This knowledge was used to understand the natural world and transfer the experiences in personal, social and global context. In the framework, instead of "sciences", "systems" preferred in describing the major fields in PISA with the aim of showing relationships between categories. The four categories in knowledge of science were mainly "Physical Systems", Living Systems", "Earth and Space Systems" and "Technology Systems".

In TIMSS 2007, there were four major content domains defining science content in the eighth grade. These were biology, chemistry, physics and Earth science. These content areas were not taught in science lessons in every participating country. In some of the countries, these subtopics were taught in geography, health education or social studies.

The knowledge of science in PISA 2006 and the content of TIMSS 2007 are compared in the Table 4.1.

Table 4.1 The Content Comparison of TIMSS 2007 and PISA 2006

TIMSS 2007	PISA 2006
Biology	Living Systems
Characteristics, classification, and life processes of organisms Cells and their functions Life cycles, reproduction, and heredity	Cells (*e.g.* structures and function, DNA, plant and animal)
Diversity, adaptation, and natural selection	Populations (*e.g.* species, evolution, biodiversity, genetic variation)
Ecosystems	Ecosystems (*e.g.* food chains, matter, and energy flow) Biosphere (*e.g.* ecosystem services, sustainability)
Human health	Humans (*e.g.* health, nutrition, disease, reproduction, subsystems [such as digestion, respiration, circulation, excretion, and their relationship])
Physics	Physical Systems
Physical states and changes in matter	Properties of matter (*e.g.* changes of state, thermal and electrical conductivity)
Energy transformations, heat, and temperature	Energy and its transformation (*e.g.* conservation, dissipation, chemical reactions)

Table 4.1 (continued) The Content Comparison of TIMSS 2007 and PISA 2006

TIMSS 2007	PISA 2006
Physics	Physical Systems
Light	Interactions of energy and matter (*e.g.* light and radio waves, sound and seismic waves)
Sound	Interactions of energy and matter (*e.g.* light and radio waves, sound and seismic waves)
Electricity and magnetism	
Forces and motion	Motions and forces (*e.g.* velocity, friction)
Chemistry	Physical Systems
Classification and composition of matter	Structure of matter (*e.g.* particle model, bonds)
Properties of matter	
Chemical change	Chemical changes of matter (*e.g.* reactions, energy transfer, acids/bases)
Earth Science	Earth and space systems
Earth's structure and physical features	Structures of the Earth systems (*e.g.* lithosphere, atmosphere, hydrosphere)
	Earth's history (*e.g.* fossils, origin and evolution)
Earth's processes, cycles, and history	Change in Earth systems (*e.g.* plate tectonics, geochemical cycles, constructive and destructive forces)

TIMSS 2007	PISA 2006
Earth Science	Earth and space systems
Earth's resources, their use and conservation	Energy in the Earth systems (*e.g.* sources, global climate)
Earth in the solar system and the universe	Earth in space (*e.g.* gravity, solar systems)
	"Technology systems"
	Role of science-based technology (*e.g.* solve problems, help humans meet needs and wants, design and conduct investigations)
	Relationships between science and technology (*e.g.* technologies contribute to scientific advancement)
	Concepts (*e.g.* optimisation, trade-offs, cost, risk, benefit)
	Important principles (*e.g.* criteria, constraints, cost, innovation, invention, problem solving)

When TIMSS and PISA was compared it was seen that in TIMSS, there wasn't such a subscale like knowledge about science as in PISA. Knowledge about science was defined to understand the purposes and the nature of obtaining evidence in other words scientific enquiry and to understand the scientific explanations resulted from the scientific enquiry. Scientific explanations meant the aims of science and how a scientist uses data. Knowledge about science was given in two titles in PISA 2006 namely Scientific Enquiry and Scientific Explanation. The types of knowledge about science were given in the Table 4.2.

Table 4.2 Types of Knowledge about Science for PISA 2006

Scientific Enquiry
• Origin (e.g. curiosity, scientific questions) • Purpose (e.g. to produce evidence that helps answer scientific questions, such as current ideas, models and theories to guide enquiries) • Experiments (e.g. different questions suggest different scientific investigations, design) • Data (e.g. quantitative [measurements], qualitative [observations]) • Measurement (e.g. inherent uncertainty, replicability, variation, accuracy/precision in equipment • and procedures) • Characteristics of results (e.g. empirical, tentative, testable, falsifiable, self-correcting)
Scientific Explanations
• Types (e.g. hypothesis, theory, model, scientific law) • Formation (e.g. existing knowledge and new evidence, creativity and imagination, logic) • Rules (e.g. logically consistent, based on evidence, based on historical and current knowledge) • Outcomes (e.g. new knowledge, new methods, new technologies, new investigations)

Note. From *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006* (p.32), 2006, Paris: OECD. Copyright 2006 by OECD. Reprinted with permission.

Scientific Enquiry was a different subscale in PISA, but in TIMSS, scientific inquiry was explained as overarching strand which covered in both content and cognitive skills. In TIMSS, it was explained that there were scientific inquiry providing to understand the natural world in all science areas including both content and cognitive components. Students were expected to demonstrate the following skills of scientific inquiry aspects:

- “Formulating questions and hypotheses
- Designing investigations
- Representing data
- Analyzing and interpreting data
- Drawing conclusions and developing explanations” (TIMSS 2007 Assessment Frameworks, 2005, p. 76).

At the eighth grade, students were expected to show scientific understanding and formulate explanations in terms of cause-effect relationships, also find alternative explanations and applying them in new situations (TIMSS 2007 Assessment Framework, 2005).

Knowledge about science in PISA was similar to science process skills in some extent. Also in both PISA 2000 and PISA 2003, the competencies were called as processes. In the framework processes were defined as the mental or sometimes physical actions used for to obtain, interpret and use evidences to have the knowledge. The scientific processes were composed of when the subject matter was from daily life. In the framework of PISA 2003, it was explained that scientific processes were implying skills and understanding needed to collect and interpret evidence from daily life and to have conclusions from it. While at schools, development of the scientific process skills was taught, few students needed these skills after school life. Furthermore, it has declared that traditionally learned scientific process skills were not the way of developing scientific knowledge. In the framework, processes were explained as processes about science, not processes within science. Scientific processes within science were used when applying science (PISA, 2003). Science process skills with the explanations were listed in Table 4.3.

Table 4.3 Science Process Skills

Basic Science Process Skills	
Observation	Using the senses to collect information
Inference	According to previous gathered data, making statements to provide an explanation
Measurement	Using instruments to describe the dimensions of an object or event.
Communication	Describing an event by using words or graphical symbols
Classification	Sorting or ordering objects or events according to a criteria or property
Prediction	Guessing the outcome of a future event based on observations or a pattern of evidence.
Integrated Science Process Skills	
Controlling variables	Being able to identify variables affecting an experimental outcome
Defining operationally	Describing to measure a variable in an experiment
Formulating hypotheses	Identify the expected outcome of an experiment
Interpreting data	Organizing data and drawing conclusions from it
Experimenting	Being able to conduct an experiment and interpret the results
Formulating models	Creating a mental or physical model of a process or event

Note. Revised from *The science process skills* by Padilla, M.J. (1990). Retrieved 3rd September from http://www.narst.org/publications/research/skill.cfm

Keywords in science process skills and knowledge about science indicated these similarities. For instance, in scientific enquiry the explanation of "purpose" was as the following. Producing evidence that helps answer scientific questions such as current ideas, models and theories to guide enquiries were similar to "inference" in scientific process skills which explained as according to previous gathered data, making statements to provide an explanation. Another similarity was between "outcomes" scientific explanations and prediction in scientific process skills. Prediction was explained as guessing the outcome of a future event based on observations or a pattern of evidence and outcome was explained as new knowledge, new methods, new technologies and new investigations.

4.1.2 Competencies in PISA vs. Cognitive Skills in TIMSS

PISA and TIMSS described subscales of science with different names. While PISA named the expected behaviors in the scales as "Competencies", TIMSS used "Cognitive Domain" for the expected student behaviors. Both Competencies in PISA and Cognitive Domain in TIMSS include three subscales.

In PISA 2006, there were three competencies namely identifying scientific issues, explaining phenomena scientifically and using scientific evidence. The achievement levels in these three competencies could give the relative strengths of countries and the areas which the countries should be strengthened (PISA, 2007). Students should follow a sequence of identifying the question, explaining phenomena and using the results while solving a science problem. In traditional science teaching, key terms and phrases in other words, explaining phenomena, are more important.

PISA assessed competencies, knowledge and attitudes and these were presented in the context. In other words, it was crucial that PISA assesses competencies, knowledge and attitudes via contexts by the end of compulsory years of schooling. In PISA 2006, contexts were chosen in terms of students' interests and lives and linguistic and cultural differences of countries were taken into consideration.

The first competency in PISA 2006 was identifying scientific issues. Identifying scientific issues included recognizing issues that are possible to investigate scientifically, identifying keywords to search for scientific information and recognizing the key features of a scientific investigation (PISA, 2007). Students should have knowledge about science in addition to knowledge of science.

Identifying scientific issues items constituted 22 % of science tasks in PISA 2006. It was most applicable to "Physical Systems", "Life Systems" and "Earth and Space Systems". All the items of this subscale were in knowledge about science in PISA 2006.

The second competency in PISA 2006 was to explaining phenomena scientifically. This competency included applying knowledge of science in a given situation, describing or interpreting phenomena scientifically and predicting changes and identifying appropriate descriptions, explanations, and predictions.

Students had to apply the appropriate knowledge of science in a given situation to define or interpret phenomena scientifically and predict changes (PISA, 2007). The percentage of items with explaining phenomena scientifically was 46 % in PISA 2006. This competency was most related to traditional science courses such as Physics and Biology in general (PISA, 2007).

The third competency was using scientific evidence in PISA 2006. "Using scientific evidence" included interpreting scientific evidence and making and communicating conclusions, identifying the assumptions, evidence and reasoning behind conclusions and reflecting on the societal implications of science and technological developments (PISA, 2007). Using scientific evidence was explained as interpreting the evidence in order to reach a decision, identify the assumptions, evidence and clarifying the reasoning. In order to use scientific evidence, students had to synthesize knowledge of science and knowledge about science while applying this knowledge in life situation. Moreover, using scientific evidence included of giving opinions for or against in a given situation based on scientific evidences in other words, concluding about scientific evidences and choosing the conclusion from alternatives in the light of scientific evidences. Students had to make logical connections between evidences and conclusions. In PISA 2006, using scientific evidence competency was questioned in 32 % of the science tasks.

In addition to intersections between knowledge about science and Science Process Skills, there were also similarities between the competencies of PISA 2006 with science process skills in Table 4.4. While students were identifying scientific issues, they were using their senses to collect information like in observation in science process skills. Another intersection was between explaining phenomena scientifically and inference because in explaining, students were using gathered data to provide an explanation. Also when using scientific evidence, students were using interpreting data and so drew conclusions from data.

Table 4.4 The Competencies in PISA 2006 vs. Science Process Skills

Competencies in PISA 2006	Related Science Process Skills
Identifying scientific issues	Observation
Explaining phenomena scientifically	Inference
Using scientific evidence	Formulating hypothesis Interpreting data

In TIMSS 2007, there were three subscales in the Cognitive Domain namely knowing, applying and reasoning. These subscales were configured according to what students should know while solving science questions in TIMSS 2007.

Knowing in TIMSS 2007 included the basic knowledge of science facts, information, tools and procedures. This knowledge helped in handling with more complex cognitive activities. The science items with knowing competency constituted 30 % of science items in TIMSS 2007. In applying, students applied knowledge and theories in some of items in TIMSS 2007. The percentage of the items including applying was 35 % in TIMSS 2007. Students engaged more complex tasks in scientific reasoning in TIMSS 2007. Thirty-five percent of the science items were about reasoning (TIMSS, 2008). According to Table 4.5, there were many intersecting skills between two international studies.

Table 4.5 The Competencies in PISA 2006 and Cognitive Skills in TIMSS 2007

Science Competencies in PISA 2006	Cognitive Skills in TIMSS 2007
Identifying scientific issues	Knowing
Recognizing issues that are possible to investigate scientifically Identifying keywords to search for scientific information Recognizing the key features of a scientific investigation	Recall / Recognize Define Describe Illustrate with examples Use tools and Procedures
Explaining phenomena scientifically	Applying
Applying knowledge of science in a given situation Describing or interpreting phenomena scientifically and predicting changes Identifying appropriate descriptions, explanations, and predictions	Compare / Contrast / Classify Use Models Relate Interpret Information Find Solutions Explain
Science Competencies in PISA 2006	Cognitive Skills in TIMSS 2007
Using scientific evidence	Reasoning
Interpreting scientific evidence and making and communicating conclusions Identifying the assumptions, evidence and reasoning behind conclusions Reflecting on the societal implications of science and technological developments	Analyze / Solve problems Integrate / Synthesize Hypothesize / Predict Design / Plan Draw conclusions Generalize Evaluate Justify

In PISA 2006 and TIMSS 2007, there were common keywords in the structure of the subscales. In identifying scientific issues and knowing, recognizing was the common keyword. Furthermore, in Applying subscale of TIMSS the behavior of “Explaining” was similar to “Identifying appropriate descriptions, explanations and predictions” in “explaining of phenomena scientifically”. In “Explaining”, students demonstrated understanding of science concept by providing an explanation for an observation or natural phenomena.

Another example of similarity was between “Drawing conclusions” in “Reasoning” and “Interpreting scientific evidence and making and communicating conclusions” in the competency of “Using scientific evidence” in PISA. “Drawing conclusions” in TIMSS 2007 was explained as to describe data trends, interpolate or extrapolate from data or given information, to make inferences in the light of scientific evidences and in order to find some logical conclusions from the questions or hypothesis. “Generalizing” was a similar behavior with “Interpreting scientific evidence and making and communicating conclusions” in the competency of “Using scientific evidence”. “Generalizing” meant to make general conclusions in given conditions, to use conclusions in new situations or to find general formulas showing physical relationships in “Reasoning” in TIMSS 2007.

4.2 Analyzing Items in PISA and TIMSS

Several similarities were found from the keywords in the structure of each subscale by Content Analysis but it was also important to analyze the items to find out differences in the item structure for each subscale. The items in PISA and TIMSS were analyzed and found some common items.

4.2.1 Identifying Scientific Issues in PISA 2006 vs. Knowing in TIMSS 2007

Although the keywords in identifying scientific issues and knowing were similar, when the items were analyzed it was found that there were differences in two subscales. As seen in released item for identifying scientific issues, the items were part of scientific investigations. It was also indicated in PISA 2006 the items in identifying scientific issues were part of scientific inquiry from knowledge about science.

In PISA, most of the released questions from identifying scientific issues were about testing the hypothesis with scientific investigation and there were also items about the nature of the scientific investigation. For instance, the aim of using the control groups or substances in an experiment was questioned in PISA. The example of identifying scientific issues was given in Item 1 while the item 2 was selected as an example to knowing. The reason for choosing item 2 was that the representativeness of other items in knowing.

Item 1

The assessment : PISA 2006

Knowledge : Knowledge about science, Scientific Enquiry

Competency : Identifying scientific issues / Recognizing the key features of a scientific investigation

SUNSCREENS

Mimi and Dean wondered which sunscreen product provides the best protection for their skin. Sunscreen products have a Sun Protection Factor (SPF) that shows how well each product absorbs the ultraviolet radiation component of sunlight. A high SPF sunscreen protects skin for longer than a low SPF sunscreen. Mimi thought of a way to compare some different sunscreen products. She and Dean collected the following:

- *two sheets of clear plastic that do not absorb sunlight;*
- *one sheet of light-sensitive paper;*
- *mineral oil (M) and a cream containing zinc oxide (ZnO); and*
- *four different sunscreens that they called S1, S2, S3, and S4.*

Mimi and Dean included mineral oil because it lets most of the sunlight through, And zinc oxide because it lmost completely blocks sunlight.

Dean placed a drop of each substance inside a circle marked on one sheet of plastic, then put the second plastic sheet over the top. He placed a large book on top of both sheets and pressed down.

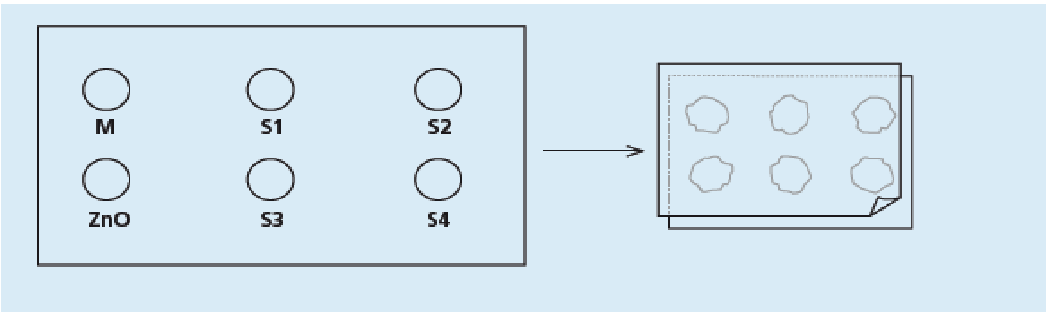
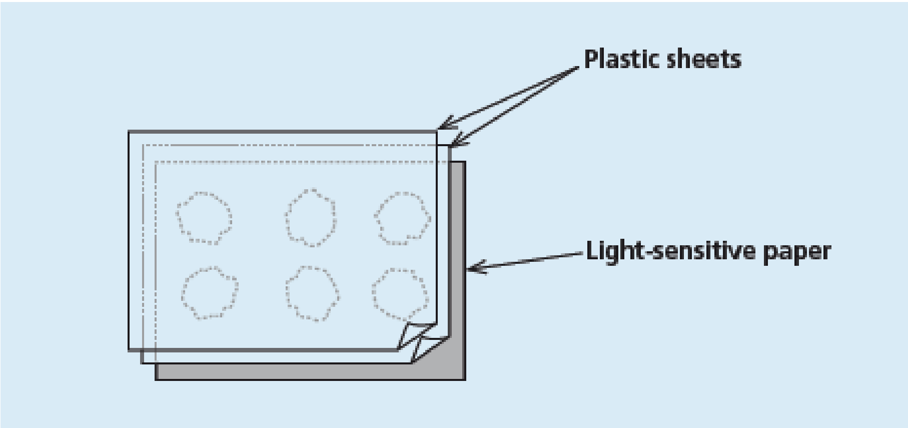


Figure 4.1 The Item 1 Representing Identifying Scientific Issues in PISA 2006

Mimi then put the plastic sheets on top of the sheet of light-sensitive paper. Light-sensitive paper changes from dark grey to white (or very light grey), depending on how long it is exposed to sunlight. Finally, Dean placed the sheets in a sunny place.



Why was the second sheet of plastic pressed down?

A. To stop the drops from drying out.

B. To spread the drops out as far as possible.

C. To keep the drops inside the marked circles.

D. To make the drops the same thickness.

Figure 4.1 (continued) The Item 1 Representing Identifying Scientific Issues in PISA 2006

Item 2

The assessment : TIMSS 2007

Content Domain : Chemistry

Cognitive Domain : Knowing / Recall /Recognize

Which of the following is closest to the percentage of the total water on earth that is fresh water?

A) 100 %

B) 90 %

C) 70 %

D) 3 %

Figure 4.2 The Item 2 Representing Knowing in TIMSS 2007

In PISA, the correct answer of the Item 1 was D and students were expected to recognize to control a variable in a scientific enquiry and understand the need of the same thickness of sunscreens. The question belonging to knowing subscale in TIMSS 2007 assesses the students' recall of a scientific knowledge.

Although the keywords in PISA and TIMSS for identifying scientific issues (ISI) and knowing seemed similar, subscales were questioning different skills as seen from the items. The percentages of correct answers also showed that there was a difference in two subscales. While the correct answers and incorrect answers were nearly the same for the Item 2 in knowing subscale in TIMSS 2007, in Item 1 the incorrect answers had a high percentage with 67.8 % in PISA 2006 as seen in Table 4.6.

Table 4.6 The Percentage of Answers for ISI and Knowing

ITEM	Frequency of Correct Answer	Correct Answer %	Frequency of Incorrect Answer	Incorrect Answer %
Item 1 (PISA)	480	32.8	1013	67.8
Item 2 (TIMSS)	292	48	319	52

4.2.2 Explaining Phenomena Scientifically in PISA 2006 vs. Applying in TIMSS 2007

Item 3

The assessment : PISA 2006

Knowledge : Knowledge of science, Physical Science

Competency : Explaining phenomena scientifically / Identifying appropriate descriptions, explanations and predictions

---

MARY MONTAGU

Read the following newspaper article and answer the questions that follow.

The History of Vaccination

Mary Montagu was a beautiful woman. She survived an attack of smallpox in 1715 but she was left covered with scars. While living in Turkey in 1717, she observed a method called inoculation that was commonly used there. This

---

Figure 4.3 The Item 3 Representing Explaining Phenomena Scientifically in PISA 2006 *treatment involved scratching a weak type of smallpox virus into the skin of healthy young people who then became sick, but in most cases only with a mild form of the disease. Mary Montagu was so convinced of the safety of these inoculations that she allowed her son and daughter to be inoculated.*

In 1796, Edward Jenner used inoculations of a related disease, cowpox, to produce antibodies against smallpox. Compared with the inoculation of smallpox, this treatment had less side effects and the treated person could not infect others. The treatment became known as vaccination.

Give one reason why it is recommended that young children and old people, in particular, should be vaccinated against influenza (flu).

Figure 4.3 (continued) The Item 3 Representing Explaining Phenomena Scientifically in PISA 2006

<u>Item 4</u>

The assessment : TIMSS 2007

Content Domain : Biology

Cognitive Domain : Applying / Explain

Keith had influenza. He played a game with two friends. One of his friends caught influenza, but the other friend did not.

What could have been the reason why one of the friends did NOT catch influenza?

Figure 4.4 The Item 4 Representing Applying in TIMSS 2007

In PISA 2006, students were expected to explain the reason as weaker immune system of children and old people in the Item 3. TIMSS 2007 asked the item 4 in a different way but the answer was the same. In TIMSS 2007, students were expected to explain the reason as the difference in immune systems of two children which one of them had a weaker immune system. In both of the questions there was a comparison in fact. While in PISA the comparison was between children and old people with others having stronger immune system, in TIMSS, the comparison was between two children which one of them had stronger immune system. To find out if there was a difference in two subscales, the percentages of correct answers of Turkish students were compared in Table 4.7.

Table 4.7 The Percentage of Answers for EPS and Applying

ITEM	Frequency of Correct Answer	Correct Answer %	Frequency of Incorrect Answer	Incorrect Answer %
Item 3 (PISA)	844	56.5	650	43.5
Item 4 (TIMSS)	438	75.8	140	24.2

Applying in TIMSS and explaining phenomena scientifically in PISA were found similar in terms of the structure of subscales and also with the item type. It was also necessary to compare the percentages of correct answers. In both PISA and TIMSS, the correct answers were more than incorrect answers but it was found that there was more difference between the correct and incorrect answers in TIMSS. The reason of obtaining higher score in PISA, could be the clue of different resistance of groups in fact but this clue could be found with being good at reading skills.

4.2.3 Using Scientific Evidence in PISA 2006 vs. Reasoning in TIMSS 2007

Item 5

The assessment : PISA 2006

Knowledge : Knowledge of science, Earth and Space Systems

Competency : Using scientific evidence / Identifying the assumptions, evidence and reasoning behind conclusions

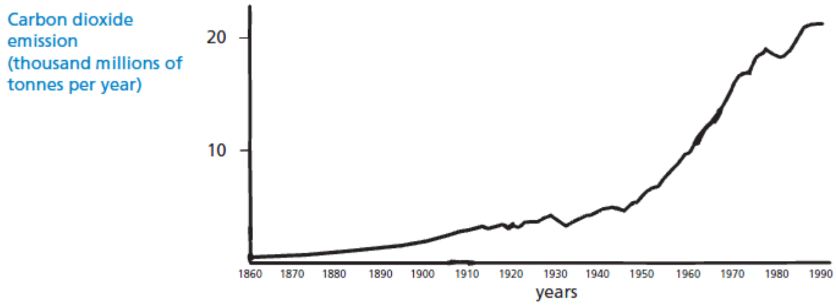
---

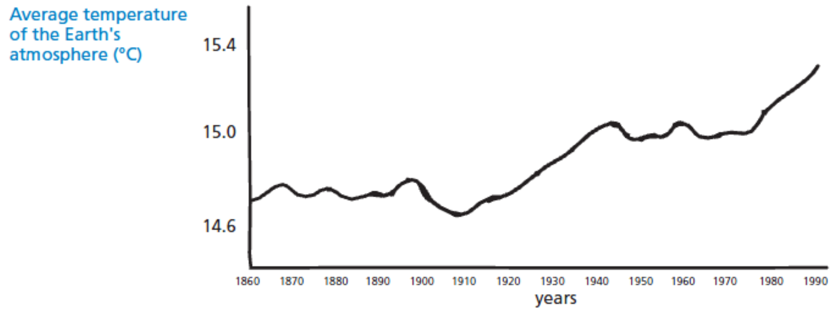
GREENHOUSE

A student named André becomes interested in the possible relationship between the average temperature of the Earth's atmosphere and the carbon dioxide emission on the Earth.

In a library he comes across the following two graphs.

Figure 4.5 The Item 5 Representing Using Scientific Issues in PISA 2006





André concludes from these two graphs that it is certain the increase in the average temperature of the Earth's atmosphere is due to the increase in the carbon dioxide emission on Earth.

What is it about the graphs that supports André's conclusion?

Figure 4.5 (continued) The Item 5 Representing Using Scientific Issues in PISA 2006

<u>Item 6</u>

The assessment : TIMSS 2007

Content Domain : Earth Science

Cognitive Domain : Reasoning / Evaluate

A plot of land was divided into 10 equal areas. A different amount of fertilizer was added to each area. Rice was planted in each area. The table below shows the amount of fertilizer added and the yields of rice for each area.

	Area									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Amount of fertilizer added (Units of Nitrogen per area)	*0*	*30*	*50*	*60*	*70*	*80*	*100*	*120*	*140*	*160*
Yield of rice (kg of rice per area)	*7.1*	*8.3*	*14.2*	*25.4*	*26.2*	*26.2*	*26.2*	*26.1*	*17.6*	*14.4*

Look at the data in the table. Provide an explanation for the effect of the amount of fertilizer on the yield of rice.

Figure 4.6 The Item 6 Representing Applying in TIMSS 2007

In the question of PISA, students were expected to compare and interpret the graphs in Item 5. Then, they would use the graphical data sets in supporting conclusion. Similarly in TIMSS, students were expected to interpret the data in the table and to use the datasets to reach a conclusion in Item 6.

In both questions in PISA and TIMSS, the information was given in a chart or a graph and students were expected to interpret them. It was also helpful to compare the percentages of correct answers of Turkish students in PISA and TIMSS in terms of selected items.

Table 4.8 The Percentage of Answers for USE and Reasoning

ITEM	Frequency of Correct Answer	Correct Answer %	Frequency of Incorrect Answer	Incorrect Answer %
Item 5 (PISA)	41.5	628	58.5	886
Item 6 (TIMSS)	42.6	202	57.4	272

The Table 4.8 showed that the percentages of Item 5 and Item 6 were very similar to each other. Therefore, it could be said there were similarities between reasoning in TIMSS and using scientific evidence in PISA 2006. Also, the evidences from the structure of subscales indicated that there were similarities between reasoning in TIMSS and using scientific evidence in PISA.

CHAPTER 5

CONCLUSIONS, DISCUSSIONS AND IMPLICATIONS

This chapter included the summary of the study, conclusions and discussion of the results and implications and recommendations for further studies.

5.1 Summary of the Study

This study investigated framework compatibility of PISA 2006 and TIMSS 2007 in science. Moreover, the similarities and differences between subscales of PISA and TIMSS were explored in this study.

5.2. Conclusions and Discussions

The framework comparison of PISA 2006 and TIMSS 2007 was investigated in two parts. The first part included content comparison in which PISA referred as knowledge and TIMSS called as content skills. In PISA, knowledge subscale was divided into two as knowledge of science and knowledge about science. Knowledge about science in PISA 2006 was similar to science process skills in some extent. Science process skills were the steps used to conduct an investigation and conclude a decision. Two dimensions of knowledge about science were given as Scientific Enquiry and Scientific Explanation. Knowledge about science assessed students' understanding purposes, nature of scientific enquiry and the results of scientific enquiry in other words scientific explanations.

Content comparison revealed that in TIMSS, the subjects covered were more than PISA for the subscales of Biology, Physics and Chemistry. Another difference was that in PISA there was an extra subscale as technology systems. In technology systems, it was emphasized more daily life applications of science in addition to technology and science relationship. Also, in this subscale as the important principles, students were encouraged to make innovations or inventions. In other words, this was a subscale which Turkish students were not accustomed to. It was also important to indicate that only eight percent of the science questions were from technology systems.

The subscales of PISA and TIMSS had several similarities in terms of topics covered in assessments except technology systems in PISA. In TIMSS 2007, Biology subscale was similar to Living Systems in PISA 2006. In PISA, with the name of Physical Systems, the subjects of Chemistry and Physics were combined. In TIMSS, characteristics, classification, and life processes of organisms, Life cycles, reproduction, and heredity for biology; electricity and magnetism for physics and properties of matter for chemistry were different subjects which was not in PISA 2006. Although the following names of the subscales differed, the subjects in Earth science in TIMSS 2007 and Earth and space systems had the same topics.

When the topics were compared for each subscale in PISA and TIMSS, most of the topics were found common in both studies but the origin of the items was totally different. In PISA, items were extracted from news, scientific articles in other words from daily life while in TIMSS, the items were more curriculum-based. Items were controlled if they were consistent with the curriculums of the countries in TIMSS.

The second part in the content analysis was the analysis of competencies in PISA and cognitive skills in TIMSS. In PISA, the competencies were more part of a scientific investigation. Students were expected at first identify scientific issues like observing in science process skills because they were encouraged to find scientific issues or investigation around. Then, students would explain phenomena scientifically like inference and communicating in science process skills. Then, students were expected to use the scientific evidence and conclude a decision by formulating hypothesis and interpreting data like in science process skills. Although there were similarities between science process skills and competencies in terms of keywords, several studies explained these process skills in PISA were not the same with the traditional scientific process skills (Olsen, 2004; PISA, 2003). Also, the competencies were called as scientific processes in the frameworks of PISA 2000 and 2003. The frameworks emphasized that the scientific processes in PISA were not scientific processes within science in other words, they weren't used while applying science, they were about scientific process skills. They also emphasized that processes could not be a content free process. The subject matters were from daily life. Therefore, it can be said that scientific processes or scientific competencies were measuring students' skills not applying those in science but in life.

Similarly, there were interceptions between cognitive skills of TIMSS and Bloom's Taxonomy of Cognitive Domain. The cognitive skills in TIMSS 2007 have common skills with Cognitive Skills of Bloom's Taxonomy. As seen in Table 5.1, the main two cognitive skills in TIMSS 2007 were very similar to the Bloom's Taxonomy.

Table 5.1 Cognitive Domain in Bloom's Taxonomy vs. Cognitive Skills in TIMSS 2007

Cognitive Domain in Blooms Taxonomy	Cognitive Skills in TIMSS 2007
Knowledge	Knowing (Define, Describe)
Comprehension	Compare / Contrast / Classify
Application	Applying
Analysis	Analyze / Solve problems
Synthesis	Integrate / Synthesize
Evaluation	Evaluate

Other subscales like "analyze/solve problem", "integrate/synthesize" or "evaluate" had the same name with cognitive domain of Bloom's Taxonomy. In

TIMSS 2007, the cognitive skills were explained in detail. For instance, in "Recalling/ Recognizing", students were expected of identifying the attributes of specific organisms and processes in addition to making the correct statements about science facts, concepts, relationships and processes. Another behavior "describing" was explained as to describe physical materials, organisms, science processes showing knowledge of structure, function and relationships. In applying subscale, "Relating" meant to relate knowledge of a physical or biological concept to an observed behavior or organisms, materials. In reasoning subscale "Evaluating" referred to comprehend the advantages and disadvantages in order to conclude about alternative processes, materials, to evaluate the effect of science and technology on biological and physical systems in the light of scientific and social factors and to evaluate results of investigations to have a conclusion with the given data.

While the subscales of scientific competencies and cognitive skills were analyzed on surface, it was found that "identifying scientific issues" was similar with knowing, "explaining phenomena scientifically" was similar with applying and "using scientific evidence" was similar with reasoning. The similarity of competencies between cognitive skills was not only observed for the common key words for each assessment, but also the general meaning of the subscales. For instance, in using scientific evidence in PISA, the competencies were about formulating hypothesis, concluding a decision and reasoning behind conclusion. These were similar with hypothesizing/predicting, drawing conclusions and justifying in reasoning in TIMSS 2007. In the same way, students were relating topics to their own knowledge, interpreted the given data and found solutions to problems, then they explained the results in TIMSS. These steps in applying in TIMSS were similar to the ones in explaining phenomena scientifically in PISA which students also applied the knowledge of science, interpreted the problem scientifically and then identified appropriate explanations.

The results of content analysis were compared with the item analysis in the current study. The items were selected in terms of the similarity of assessed cognitive skill from PISA and TIMSS. It was found that identifying scientific issues in PISA and knowing in TIMSS were different contrary to the results of content analysis. Identifying scientific issues items were part of knowledge about science which was more part of scientific investigation. The items of this subscale assessed to find out the possible scientific investigations and understand the key features of scientific investigation. Whereas knowing in TIMSS assessed students' recall of scientific knowledge and describing scientific processes. The difference between subscales of PISA and TIMSS were also supported with the percentages of items. Explaining scientific phenomena in PISA and applying in TIMSS were found similar in terms of item analysis. In both selected items from PISA and TIMSS, students were expected to explain the difference in the immunity system of people. Although the items were open response and the keywords in the structure of subscales were similar, the percentages of correct answers of Turkish students differed. The reason might be the deficiencies of Turkish students in reading skills because there was a passage in PISA. This was supported with the study of Ruddock et al. (2006). In the study, it was found that the grammatical complexity in science items was found to be higher in PISA than in TIMSS (Ruddock et al., 2006). Moreover, this situation may result from Turkish students were not good at transferring knowledge to daily life. Although they knew immune system of people was different for every individual, they could not be successful when it was asked in a way using the knowledge in daily life. It was also significant to say that the percentages of correct answers were still more than incorrect answers in both assessments for the selected items. The third comparison between subscales was using scientific evidence in PISA to reasoning in TIMSS. The items from PISA and TIMSS were assessing to interpret the graphs or table which was both open response. The keywords in the structures of subscales and the similar items showed that using scientific evidence in PISA and reasoning were similar. Another proof of similarity was the percentages of correct answers of Turkish students for the selected items. The percentages of two subscales were nearly the same.

The item analysis showed that the competencies in PISA and the cognitive skills in TIMSS had similarities in addition to differences. This difference came from identifying scientific issues in PISA which was more part of a scientific investigation. There were studies asserting that there cannot be direct comparisons between PISA and TIMSS (Adams, 2003; Harding, 2003; Olsen, 2004). The reason was explained as differences in target population, assessing aims and measured scientific concepts. Adams (2003) declared that the assessments were not statistically linked. There were also the studies comparing two assessments in the item level (Dossey, 2006; Nohara, 2001; Ruddock et al; 2006; Wu, 2008) but except Wu (2008) the other authors studied the item analysis of two studies to find out some common similarities such as appropriateness to real life, familiarity to students or emphasized thinking skills. On the other hand, Wu (2008) investigated the strengths and weaknesses of 22 countries by using regression techniques among PISA and TIMSS.

In this study, similar to most studies, the scores in two assessments could not be compared because of the different target population. PISA and TIMSS were compared in a general way as in most of the studies (Bybee, 2007; Geske & Kangro, 2002; Ginsburg et al., 2005; Harlen, 2001; Hutchison & Schagen, 2007; Lie, 2005). The results of content analysis were consistent with the current study which was found that the aims, the population, sampling and assessment cycle was differed. Therefore, the scores of students could not be compared in PISA and TIMSS. Moreover, it was found that there were similar items in PISA and TIMSS assessing the same skills and so only the percentages of the items could be compared in PISA and TIMSS.

5.3 Summary of Discussion

The research questions were:

1) Are the PISA and TIMSS comparable?

2) Are the science frameworks of TIMSS and PISA parallel?

It was found that PISA and TIMSS could not be compared statistically because of the different populations but with the common items in each assessment, the percentage of correct answers could be compared. With this comparison, it was investigated that the science frameworks of PISA and TIMSS were similar in some extent. While identifying scientific issues was different from knowing, explaining phenomena scientifically in PISA vs. applying in TIMSS and using scientific evidence in PISA vs. reasoning in TIMSS were found similar. It was also important to indicate that the percentages of correct answers in explaining phenomena scientifically were lower than applying for the selected items. The reason was explained as follows. The reading ability and transferring the knowledge to daily life were also assessed in PISA. Furthermore, the complexity of grammar in PISA made difficult to have better results for Turkish students.

5.4 Implications of the Study

The results of the present study have some implications for governors, teachers and researchers.

The present study indicated that Turkish students were not successful in transferring knowledge to daily life. This stemmed from the old curriculum which was based on traditional teaching. The results of the study may be used in using different teaching methods emphasizing using the knowledge in daily life. For instance, teachers may use inquiry based learning or do experiments related with daily life.

Not only the teachers, but also governors may use the results of this study. Governors may change the curriculum that students can relate their scientific knowledge to daily life. New science curriculum was based on constructivism but the exams were still forcing students to memorize knowledge rather than thinking. In order to prevent rote memorization, not only the curriculum but also the national exams should be changed because if in national exams, there were questions encouraging students to memorize, then the current curriculum change would not work. By this way, Turkish students could also be successful in the life-related questions.

5.5 Recommendations for Further Research

Present study has suggested some useful research topics:

1. The common items can be selected with the idea of experts. By analyzing more items, the results of comparisons would be more comprehensive.

2. The study included the 15 year-old students for PISA and eighth grade students for TIMSS. These students were taught with old curriculum. Therefore, as a comparison, researchers may study PISA with the students taught with new curriculum. Therefore, the effectiveness of the constructivist curriculum might be questioned in this way.

3. The similar item analysis and comparison of PISA and TIMSS can be done for Mathematics in 2012 which was the major domain was again Mathematics.

REFERENCES:

Acar, O. (2008). PISA Sonuçları Işığında Türkiye'nin Rekabet Gücünün Değerlendirilmesi. Retrieved August 20, 2010 from http://www.tepav.org.tr/tur/admin/dosyabul/upload/Politika_notu_ozan_acar.pdf

Adams, R. J. (2003). Response to 'Cautions on OECD's Recent Educational Survey (PISA)'. *Oxford Review of Education*, *29*(3), 377-389.

Australian Council for Educational Research. (2010). *OECD Programme for International Student Assessment (PISA)*. Retrieved July 25, 2010 from http://www.acer.edu.au/ozpisa/

Beaton, A. E., Postlethwaite, T. N., Ross, K. N., Spearritt D., & Wolf R. M. (1999). *The Benefits and Limitations of International Educational Achievement Studies.* Paris: International Institute for Educational Planning/International Academy of Education.

Bybee, R. W. (2007). Science teaching and international assessments: An introduction to PISA and TIMSS. *The Science Teacher, 74*, 41-48.

Ciascai, L. (2009). Comparative study on Romanian school science curricula and the curriculum of TIMSS 2007 testing. *Acta Didactica Napocensia, 2*(2), 25-34.

Creemers, B. P. M. (2006). The importance and perspectives of international studies in educational effectiveness. *Educational Research and Evaluation, 12*(6), 499 – 511.

Çalışkan, M., Alkan, M., Taşkın, T., Panal, A., and Ovayolu, Ö. (2010). *PISA 2006 Projesi: Ulusal Nihai Rapor.* Ankara: Milli Eğitim Bakanlığı-Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı.

Dossey, J. , McCrone, S., and O'Sullivan, C. (2006). *Problem solving in international comparative assessments.* Washington, DC: National Center for Education Statistics.

Geske, A., and Kangro, A. (2002). Evaluation of Latvia's science education in the IEA TIMSS and OECD PISA framework. *Journal of Baltic Science Education, 2,* 59-66.

Ginsburg, A., Cooke, G., Leinwand, S., Noell, J., and Pollock, E. (2005). *Reassessing U. S. International Mathematics Performance: New Findings from the 2003 TIMSS and PISA.* Washington, DC: American Institutes for Research.

Harding, D. (2003). *Monitoring National Levels of Student Achievement*. Retrieved March 1tst, 2010, from http://info.worldbank.org/etools/library/latestversion_p.asp?objectID=242910&lprogram=2

Harlen, W. (2001). The assessment of scientific literacy in the OECD/PISA project. *Studies in Science Education, 36,* 79 –104.

Hutchison, D., and Schagen, I. (2007). Comparisons between PISA and TIMSS - are we the man with two watches?. In Loveless, T. (Eds.), *Lessons Learned: What International Assessments Tell Us About Math Achievement.* Washington, D.C.: Brookings Institute Press.

Kjaernsli, M. (2003). *Achievement in Scientific Literacy in PISA: Content knowledge and process skills.* Retrieved August 13, 2009 from http://www1.phys.uu.nl/esera2003/programme/pdf/260S.pdf

Kjaernsli, M., and Lie, S. (2004). PISA and scientific literacy: Similarities and differences between the Nordic countries. *Scandinavian Journal of Educational Research, 48*(3), 271-286.

Lie, S. (in press). How can large international comparative studies contribute to the quality of science education? *Research and the Quality of Science Education.*

Martin, M.O., Mullis, I.V.S., and Foy, P. (2008). *TIMSS 2007 International Science Report. Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eight Grades.* Chestnut Hill, MA: IEA TIMSS & PIRLS International Study Center.

Ministry of National Education -Educational Research and Development Department (2007). *PISA 2006 Uluslararası Öğrenci Değerlendirme Programı: Ulusal Ön Rapor.* Ankara: Milli Eğitim Bakanlığı-Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı.

MONE (2004). *MEB Müfredat Geliştirme Süreci*. Retrieved August 25, 2009 from http://ttkb.meb.gov.tr/programlar/prog_giris/prog_giris_1.html

Nohara, D. (2001). *A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)* (NCES Working Paper No. 2001-07). Washington, DC: National Center for Education Statistics.

Olsen, R. V. (2004). *The OECD PISA assessment of scientific literacy: how can it contribute to science education research?* Paper presented at the National Association for Research in Science Teaching (NARST), 1.-3. April, Vancouver.

Olsen, R. V. (2005). *Achievement tests from an item perspective. An exploration of single item data from the PISA and TIMSS studies, and how such data can inform us about students' knowledge and thinking in science.* Oslo: Unipub forlag.

Padilla, M.J. (1990). The science process skills. *Research Matters – to the Science Teacher, No. 9004*. Retrieved September 3, from http://www.narst.org/publications/research/skill.cfm

Programme for International Student Assessment (PISA) (2006). *Assessing Scientific, Reading and Mathematical Literacy: A framework for PISA 2006.* Paris: OECD.

Programme for International Student Assessment (PISA) (2007). *Science Competencies for Tomorrow's World. Volume I: Analysis: PISA 2006.* Paris: OECD.

Programme for International Student Assessment (PISA) (2005). *PISA 2003 Technical Report.* Paris: OECD.

Programme for International Student Assessment (PISA) (2009). *PISA 2006 Technical Report.* Paris: OECD.

Robitaille, D.F. and Robeck, E.C. (1996). The character and the context of TIMSS. In Robitaille, D.F. and Garden, R.A. (Eds.), *Research questions & study design* (TIMSS Monograph No.2). Vancouver: Pacific Educational Press.

Ruddock, G., Clausen-May, T., Purple, C., and Ager, R. (2006). *Validation Study of the PISA 2000, PISA 2003 and TIMSS 2003 International Studies of Pupil Attainment* (DfES Research Report 772). London: DfES.

Trends in Mathematics and Science Assessments (TIMSS) (2005). *TIMSS 2007 Assessment Framework.* Chestnut Hill, MA: Boston College.

Trends in Mathematics and Science Assessments (TIMSS) (2008). *TIMSS 2007 Technical Report.* Chestnut Hill, MA: Boston College.

Turmo, A. (in press). Scientific Literacy and Learning Strategies - Empirical Results from the PISA Study. I: **Science and Technology Education for a Diverse World- Dilemmas, Needs and Partnerships.** Lublin, Polen: Marie Curie Sklodowska University Press 2006.

White, P., and Smith, E. (2005). What can PISA tell us about teacher shortages? *European Journal of Education, 40*(1), 92-112.

Wu, M.L. (2006). A comparison of mathematics performance between east and west: What PISA and TIMSS tell us. In F. Leung., K.D. Graf., & F. Lopez-Rea (Eds.), Mathematics Education in Different cultural Traditions - A Comparative study of east Asia and the west, pp 239-259. USA: Springer.

Wu, M.L. (2008). A Comparison of PISA and TIMSS 2003 achievement results in Mathematics and Science. Paper presented at the Third IEA Research Conference, Taipei, September, 2008.