

EURASIAN JOURNAL OF EDUCATIONAL RESEARCH
E Ğ İ T İ M A R A Ş T I R M A L A R I D E R Ğ İ S İ

A Quarterly Peer-Reviewed Journal, Year: 10 Issue: 40 / 2010
Üç Ayda Bir Yayınlanan Hakemli Dergi, Yıl: 10 Sayı: 40 / 2010

FOUNDING EDITOR / Kurucu Editör
Veysel Sönmez, Hacettepe University, Ankara, TURKEY

EDITORS / Editörler
Giray Berberoğlu, Middle East Technical University, Ankara, TURKEY
Ali E. Şahin, Hacettepe University, Ankara, TURKEY
Şenel Poyrazlı, Penn State University, PA, USA

INTERNATIONAL EDITORIAL BOARD / Uluslararası Editörler Kurulu
Ahmet Aypay, Eskişehir Osman Gazi University, Eskişehir, TURKEY
Anita Pipere, Daugavpils University, LATVIA
Aslı Özgün Koca, Wayne State University, USA
Beatrice Adera, Penn State University, PA, USA
Cem Birol, Near East University, Nicosia, TRNC
Danny Wyffels, KATHO University, Kortrijk, BELGIUM
Hasan Hüseyin Aksoy, Ankara University, Ankara, TURKEY
Iordanescu Eugen, Lucian Blaga University of Sibiu, ROMANIA
İsmail Ö. Zembat, United Arab Emirates University, Abu Dhabi, UAE
Mehmet Buldu, United Arab Emirates University, Abu Dhabi, UAE
Sven Persson, Malmö University, Malmö, SWEDEN
Ürsula Casanova, Arizona State University, USA
Yusif Məmmədov, Azerbaijan State Pedagogy University, Baku, AZERBIJAN

EDITORIAL OFFICE / Yayın Yönetim Yeri
Özer Daşcan: Publishing manager / Sahibi ve yazı işleri müdürü
Dilek Ertuğrul, Anı Yayıncılık, Kızıllırmak Sokak 10/A
06640 Bakanlıklar Ankara, TURKEY
ejer.editor@gmail.com
Tel: +90.312 425 81 50 pbx Fax: +90.312 425 81 11

Printing Date / Basım Tarihi: 29.07.2010
Printing Address / Matbaa Adresi: Sözkese Mat. İ.O.S. Mat. Sit. 558 Sk. No:41 Yenimahalle-Ankara
Yayın Türü: Yaygın Süreli Yayın
Cover Design / Kapak Tasarımı: Anı Yayıncılık
Typography / Dizgi: Gamze ŞAHİN
The ideas published in the journal belong to the authors.
Dergide yayımlanan yazıların tüm sorumluluğu yazarlarına aittir

Eurasian Journal of Educational Research (ISSN 1302-597X) is a quarterly peer-reviewed journal published by Anı Yayıncılık
Eğitim Araştırmaları (ISSN 1302-597X) Anı Yayıncılık tarafından yılda dört kez yayımlanan hakemli bir dergidir.
© 2010 ANI Publishing. All rights reserved. © 2010 ANI Yayıncılık. Her hakkı saklıdır.

Eurasian Journal of Educational Research (EJER) is abstracted and indexed in;
Social Science Citation Index (SSCI),
Social Scisearch,
Journal Citation Reports/ Social Sciences Editon,
Higher Education Research Data Collection (HERDC),
Educational Research Abstracts (ERA),
SCOPUS database,
EBSCO Host database, and
ULAKBİM national index.

CONTENTS
İçindekiler

Comparative Analysis of Multicriteria Decision Making Methods for Postgraduate Student Selection
Taner Altunok, Özgür Özpeynirci, Yiğit Kazançoğlu, and Recai Yılmaz 1-15

The Relationship between Demographics, Self Efficacy, and Burnout among Teachers
Nilay T. Bümen..... 17-36

Cooperative Learning Approach in the Science Teaching
Cavide Demirci 37-52

Development and Application of Three-Tier Heat and Temperature Test: Sample of Bachelor and Graduate Students
Ali Eryılmaz 53-76

Effects of Cooperative Games on Social Skill Levels and Attitudes Toward Physical Education
Osman Gülay, Dilşad Mirzeoğlu, and Müberra Çelebi 77-92

Students' Affective Characteristics and their Relation to Mathematical Literacy Measures in the Programme for International Student Assessment (PI-SA) 2003
Çiğdem İş Güzel & Giray Berberoğlu..... 93-112

Purposes and Identities of Higher Education Institutions, and Relatedly the Role of the Faculty
Zeynep Kızıltepe 113-132

A Turkish Adaptation of the Inventory of Parent and Peer Attachment: The Reliability and Validity Study
Ercan Kocayörük..... 133-151

From Interpretations to Graphical Representations: A Case Study Investigation of Covariational Reasoning
Önder Köklü & Elizabeth Jakubowski 153-172

Well-Being of University Students Who Have Lived Abroad: The Case of Hacettepe University
Hülya Şahin & Yasemin Akman Karabeyoğlu ... 173-188

Effects of Reflective Writing in Mathematics Methods Courses on Pre-Service Teachers' Perceptions
Renan Sezer 189-206

Validation of the Social Achievement Goal Orientation Scale in Iranian Students
Siavash Talepasand, Fatimah Alijani, and Iman Bigdeli 207-222

Simple Formative Assessment, High Learning Gains in College General Chemistry
Yalçın Yalaki..... 223-241

REVIEWERS of the 40th ISSUE
40. Sayı Hakemleri

Adnan Erkuş
Ali Baykal
Behiye Ubuz
Behzat Bektaşlı
Canan Candaş Karababa
Cevat Celep
Ercan Kiraz
Esin Tezer
Esra Macaroğlu Akgül
Evrin Genç Kumtepe
Fatoş Silman
Haluk Soran
Hülya Kelecioğlu
Jorge Perez
Mehmet Buldu
Mehtap Çakan
Mustafa Levent Ince
Oylum Akkuş
Prabha Betne
Rod Martin
Sekan Yılmaz
Selahattin Gelbal
Sibel Güneysu
Victor Sampson
Yeşim Bulca

Comparative Analysis of Multicriteria Decision Making Methods for Postgraduate Student Selection

Taner Altunok*
Özgür Özpeynirci**
Yiğit Kazançoğlu***
Recai Yılmaz****

Suggested Citation:

Altunok, T., Özpeynirci, Ö., Kazançoğlu, Y. & Yılmaz, R. (2010). Comparative analysis of multicriteria decision making methods for postgraduate student selection. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 40, 1-15.

Abstract

Problem Statement: The educational recourses available for postgraduate education are limited compared with the resources for undergraduate education. An academic advisor is required for each graduate student. The academic staff allocates an important portion of his/her time for graduate studies of his/her students. Since the number of academic staff is limited, the number of graduate students should be limited, too. While a large number of students may attend an undergraduate course, this may not be possible for some graduate courses, which are designed especially for small student groups. Hence, the selection of postgraduate students is an important problem. The postgraduate student selection problem can be defined as selecting a subset of students from the applicants for a postgraduate program. Naturally, this problem is a multi-criteria decision-making problem since each applicant has several attributes and these attributes should be considered simultaneously during the selection process.

* Prof.Dr., Çankaya University, Graduate School of Natural and Applied Sciences, Ankara, Turkey, taltunok@cankaya.edu.tr

** Asst.Prof.Dr., Izmir University of Economics, Dept. of Logistics Management, Izmir, Turkey, ozgur.ozpeynirci@ieu.edu.tr

*** Asst.Prof.Dr., Izmir University of Economics, Dept. of Business Administration, Izmir, Turkey, yigit.kazancoglu@ieu.edu.tr (Corresponding Author)

**** Defense Sciences Institute, Turkish Military Academy, Ankara, Turkey, ryilmaz@posta.com

Purpose of Study: The purpose of this study is to compare the performance of different multi criteria decision-making methods developed for ranking alternatives for the postgraduate student selection problem. The best method will also be used in the student selection process for the Defense Sciences Institute of Turkish Military Academy.

Methods: This paper discusses three multi-criteria decision making methods developed for ranking alternatives, namely Analytic Hierarchy Process (AHP), Weighted Product (WP) and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). A performance measure is defined and the performance values of the methods are compared using real data gathered from the graduates of Defense Sciences Institute of Turkish Military Academy for illustration purposes.

Findings and Results: According to the findings, AHP is the best of the three competitors. The ranking obtained by AHP is quite similar to the ranking of students with respect to their cumulative grade point average (CGPA) after their graduation from the program.

Conclusions and Recommendations: Considering graduate student selection problem as a multicriteria decision-making problem is very important and enables universities to enroll graduate students with important attributes and a variety of strengths. This paper compares the performance of different multicriteria decision-making methods on graduate student selection problem using a single performance measure. A further study can be making such a comparison with multiple performance measures.

Keywords: Multicriteria decision making, graduate student selection, AHP, TOPSIS.

Multicriteria Decision Making Methods

Multicriteria decision-making (MCDM) has been widely used in ranking or selecting one or more alternatives from a finite or infinite number of alternatives with respect to multiple criteria or attributes. These criteria or attributes are usually conflicting (Zeleny, 1982; Yoon & Hwang, 1995; Belton & Stewart, 2002). MCDM methods assume the existence of single decision maker (DM) or a group of DMs that is the owner of the problem. Among its broad range of real life applications, MCDM methods are used for evaluating the performance of organizations in higher education sectors in various decision contexts (e.g. Saaty & Ramarujam, 1983; Blanchard, Pierce & Hood, 1989; Davey, Olson & Wallenius, 1994; Mustafa and Goh, 1996; Yeh, 2003, Bahurmoz, 2003). In these applications, MCDM provides a systematic way for assisting the decision makers in making more informed decisions about the comparative performance of the resources and operations.

In general, MCDA problems involve six components (Keeney & Raiffa, 1976):

- A goal or a set of goals that the decision maker wants to achieve,

- The decision maker or a group of decision makers involved in the decision-making process, and their preferences with respect to the evaluation criteria,
- A set of evaluation criteria (objectives and/or physical attributes)
- The set of decision alternatives,
- The set of uncontrollable (independent) variables or states of nature (decision environment)
- The set of outcomes or consequences associated with each alternative attribute pair.

The Postgraduate Student Selection Problem Formulated as MCDM Problem

The MCDM problem involves a set of m alternatives (applicants) A_i ($i=1, 2, \dots, m$). These alternatives are rated by performance for a set of n selection criteria C_j ($j=1, 2, \dots, n$). These form a decision matrix, where rows and columns indicate m alternatives and n attributes respectively, given as

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}$$

where X_{ij} are the performance ratings of alternative A_i ($i=1, 2, \dots, m$) with respect to attribute C_j ($j=1, 2, \dots, n$).

The weight of a criterion represents its relative importance among all criteria. The weights can be calculated by various methods that gather preference information from DM. In this study we employ a pairwise comparison method, and it will be discussed in detail in the next subsection.

The objective of the problem is to rank all the alternatives in terms of their overall preference value, which is a function of the weights and the performance ratings. For PSSP, we will discuss and compare the following three methods respectively: Analytic Hierarchy Process (AHP), Weighted Product (WP) and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). Tanrıöğen and İşcan (2009) study the academic achievement of undergraduate students.

Analytic Hierarchy Process (AHP)

The AHP is designed to solve complex multi-criteria decision problems. Although it is not easy for a human to understand and solve such problems, AHP enables the division of problems into smaller pieces so that the DM can make judgments in a more reliable way. It facilitates the decision-making process by organizing perceptions, feelings, judgments and experiences into a framework. The AHP is implemented in the Expert Choice software and this software has been used in a variety of decisions and projects in more than 20 countries (Saaty, 2001).

Initially, AHP structures the problem as a hierarchy of levels. Once the hierarchy has been constructed, the decision maker begins the comparison procedure to determine the relative importance of the elements in each level. Comparison involves eliciting judgments about the dominance of one element over another with respect to each criterion.

The scale used for comparisons in AHP enables the DM to incorporate experience and knowledge intuitively and indicate how many times an element dominates another with respect to the criterion (Millet, 1997, p.1200). The DM can verbally express his/her preference between each pair of elements as equally important, moderately more important, strongly more important, very strongly more important, and extremely more important. These descriptive preferences would then be translated into numerical values 1, 3, 5, 7, and 9 respectively with 2, 4, 6, and 8 as intermediate values for comparisons between two successive qualitative judgments. Reciprocals of these values are used for the corresponding transposed judgments. The comparison scale used by AHP is given in Table 1.

Finally, all the comparisons are synthesized to rank the alternatives. The output of AHP is a prioritized ranking of the decision alternatives based on the overall preferences expressed by the decision maker. Sensitivity analysis is used to investigate the impact of changing the priorities of the criteria on the final outcome.

Table 1
The Fundamental Scale of AHP

Intensity of Importance	Definition	Explanation
1	Equal Importance	Two activities contribute equally to the objective
3	Moderate Importance	Experience and judgment slightly favor one activity over another
5	Strong Importance	Experience and judgment strongly favor one activity over another
7	Very strong Importance	An activity is favored very strongly over another; its dominance demonstrated in practice
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2,4,6,8	For compromise between the above values	Sometimes one needs to interpolate a compromise judgment numerically because there is no good word to describe.
Reciprocals of above	If activity i has one of the above nonzero numbers assigned to it when compared with activity j , then j has the reciprocal value when compared with i	A reasonable assumption

Weighted Product Method

The WP method uses the multiplication for connecting attribute ratings, each of which is raised to the power of the corresponding attribute weight (Bridgman, 1922; Starr, 1972). This multiplication process has the same effect as the normalization process for handling different measurement units. The overall preference score of each alternative (S_i) is given by

$$S_i = \prod_{j=1}^n x_{ij}^{w_j}; \quad i = 1, 2, \dots, m$$

where $\sum_{j=1}^n w_j = 1$. The weight w_j is a positive power for benefit attributes and a negative power for cost attributes. In this study, for easy comparison with other methods, the relative overall preference value of each alternative (V_i) is given by

$$V_i = \frac{\prod_{j=1}^n x_{ij}^{w_j}}{\prod_{j=1}^n (x_j^*)^{w_j}}, \quad i = 1, 2, \dots, m, \quad x_j^* = \max x_{ij} \quad 0 \leq V_i \leq 1$$

The greater the value (V_i), the more preferred the alternative (A_i).

The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

TOPSIS is based on the concept that the most preferred alternative should not only have the shortest distance from the positive ideal solution, but also have the longest distance from the negative ideal solution (Hwang and Yoon, 1981; Zeleny, 1982). The TOPSIS (technique for order performance by similarity to idea solution) was first developed by Hwang and Yoon (1981). According to this technique, the best alternative would be the one that is nearest to the positive-ideal solution and farthest from the negative ideal solution (Ertugrul & Karakasoglu, 2007). The positive-ideal solution is a solution that maximizes the benefit criteria and minimizes the cost criteria, whereas the negative ideal solution maximizes the cost criteria and minimizes the benefit criteria (Wang & Elhag, 2006). In short, the positive-ideal solution is composed of all best values attainable from the criteria, whereas the negative ideal solution consists of all worst values attainable from the criteria (Wang & Yang, 2007). There are many studies in the literature that uses TOPSIS for solving of MCDM problems. (Chu, 2002, Lai, Liu & Hwang, 1994 and Shyur, 2006). This concept has been widely used in various MCDM models for solving practical decision problems (e.g. Parkan & Wu, 1999; Deng, Yeh & Willis, 2000; Yeh, Deng & Chang, 2000). This is due to

- its simplicity and comprehensibility in concept,
- its computational efficiency,

- its ability to measure the relative performance of the decision alternatives in a simple mathematical form.

The procedure of TOPSIS can be expressed in a series of steps:

Step 1: Calculate the normalized decision matrix. The normalized value r_{ij} is calculated as

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \quad i = 1, 2, \dots, m \quad j = 1, 2, \dots, n$$

Step 2: Calculate the weighted normalized decision matrix. The weighted normalized value v_{ij} is calculated as

$$V_{ij} = w_j \cdot r_{ij} \quad i = 1, 2, \dots, m \quad j = 1, 2, \dots, n$$

where w_j is the weight of the i^{th} criterion, and $\sum_{j=1}^n w_j = 1$,

Step 3: Determine the positive ideal and negative ideal solution

$$A^+ = (V_1^+, V_2^+, \dots, V_n^+) = [V_j^+]_{1 \times n} \quad V_j^+ = \max(V_{ij}) \quad i = 1, 2, \dots, m$$

$$A^- = (V_1^-, V_2^-, \dots, V_n^-) = [V_j^-]_{1 \times n} \quad V_j^- = \min(V_{ij}) \quad i = 1, 2, \dots, m$$

Step 4: Calculate the separation measures, using the n -dimensional Euclidean distance. The separation of each alternative from the ideal solution is given as

$$S_i^+ = \sqrt{\sum_{j=1}^n (V_{ij} - V_j^+)^2} \quad i = 1, 2, \dots, m$$

Similarly, the separation from the negative ideal solution is given as

$$S_i^- = \sqrt{\sum_{j=1}^n (V_{ij} - V_j^-)^2} \quad i = 1, 2, \dots, m$$

Step 5: Calculate the relative closeness to the ideal solution. The relative closeness of the alternative A_i with respect to A^+ is defined as

$$RC_i = \frac{S_i^-}{S_i^- + S_i^+} \quad i = 1, 2, \dots, m$$

Since $S_i^- \geq 0$ and $S_i^+ \geq 0$ then, clearly, $RC_i \in [0, 1]$.

Step 6: Rank the preference order. For ranking alternatives using this index, we can rank alternatives in decreasing order.

The basic principle of the TOPSIS method is that the chosen alternative should have the "shortest distance" from the positive ideal solution and the "longest distance" from the negative ideal solution. The TOPSIS method introduces two "reference" points, but it does not consider the relative importance of the distances from these reference points.

Differences between MCDM Methods

The main differences between the three methods discussed above lie in the normalization process for comparing all performance ratings on a common scale, and the aggregation of the normalized decision matrix and weighting vector for obtaining an overall preference value for each alternative. Due to these structural differences, the ranking outcome produced by the three methods may be very different. In fact, the empirical study presented in this paper shows that the rankings are so different that the relative effectiveness of the methods used needs to be examined.

The Empirical Study

A group of academic staff is selected among the experienced members of the institute. This group of decision makers was consulted in order to determine the selection criteria and weights of these criteria. Figure 1 shows the hierarchical representation of selection criteria obtained by group decision.

Pairwise comparison was used to weight the determined selection criteria. The Expert Choice application was used for this process. Each person joining group decision makes his/her pairwise comparison independently, and the weights of criteria are determined by combining pairwise values. The weights of the main criteria are given in Table 2.

Table 2
The weights of the main criteria

Main Criteria	Weight
APPEE Grade	0.378
English Exam Grade	0.216
Interview Grade	0.173
Graduation Degree	0.111
Graduation School	0.071
Duration Since Graduation	0.051

According to Postgraduate Education Regulations, in student selection the APPEE grade weight must be at least 50% of the total weight. Therefore, we cannot employ the current weights as computed by the software. We cannot add such a constraint to the software. Hence, a reasonable approach is to set the weight of the first criterion as 0.5, remove it (and all pairwise comparisons done with this criterion) and use the software with the remaining five criteria. Note that the sum of these criteria should be normalized to 0.5, not to 1.0. However, this approach results with a loss of comparison information, which is an important disadvantage.

There are many approaches to calculating the weights from a set of pairwise comparisons. Chandran, Golden and Wasil (2005) discuss and compare these approaches. Some of these approaches use linear programming problems (Bryson, 1995, p.645). Expert Choice software, on the other hand, calculates the weights using the Eigen value approach. To the best of our knowledge, there is only one method that has the ability to handle the constraint imposed by the Postgraduate Education Regulations, Özpeynirci et al. (2009). Their approach is based on goal programming and it is possible to add a constraint that APPEE grade weighting must be at least 0.5.

The mathematical model of Özpeynirci et al. (2009) is as follows:

Notation

Parameters

j, k : criterion indices, $k=1, \dots, n$

α_{jk} = The relative importance of criterion j with respect to criterion k . (where

$$\alpha_{jk} \alpha_{kj} = 1)$$

Decision variables

w_j = The weight of criterion j .

d_{jk}^- = Negative deviation in the comparison of criteria j and k .

d_{jk}^+ = Positive deviation in the comparison of criteria j and k .

Mathematical model

$$\text{Min } \sum_{j=1}^n \sum_{k=1}^n d_{jk}^+ + \sum_{j=1}^n \sum_{k=1}^n d_{jk}^- \quad (1)$$

St.

$$w_j = \alpha_{jk} w_k + d_{jk}^+ - d_{jk}^- \text{ for all } j \text{ and } k \quad (2)$$

$$\sum_{j=1}^n w_j = 1 \quad (3)$$

$$w_1 \geq 0.5 \quad (4)$$

The objective function (1) minimizes the total amount of positive and negative deviation. All pairwise comparisons are introduced and the deviations are calculated in constraint set (2). Weights are normalized in (3). It is stated that the weight of the first criterion should be at least 0.5 in (4). Using the above mathematical model, the weights of the main criteria are computed. The resulting weights are given in Table 3.

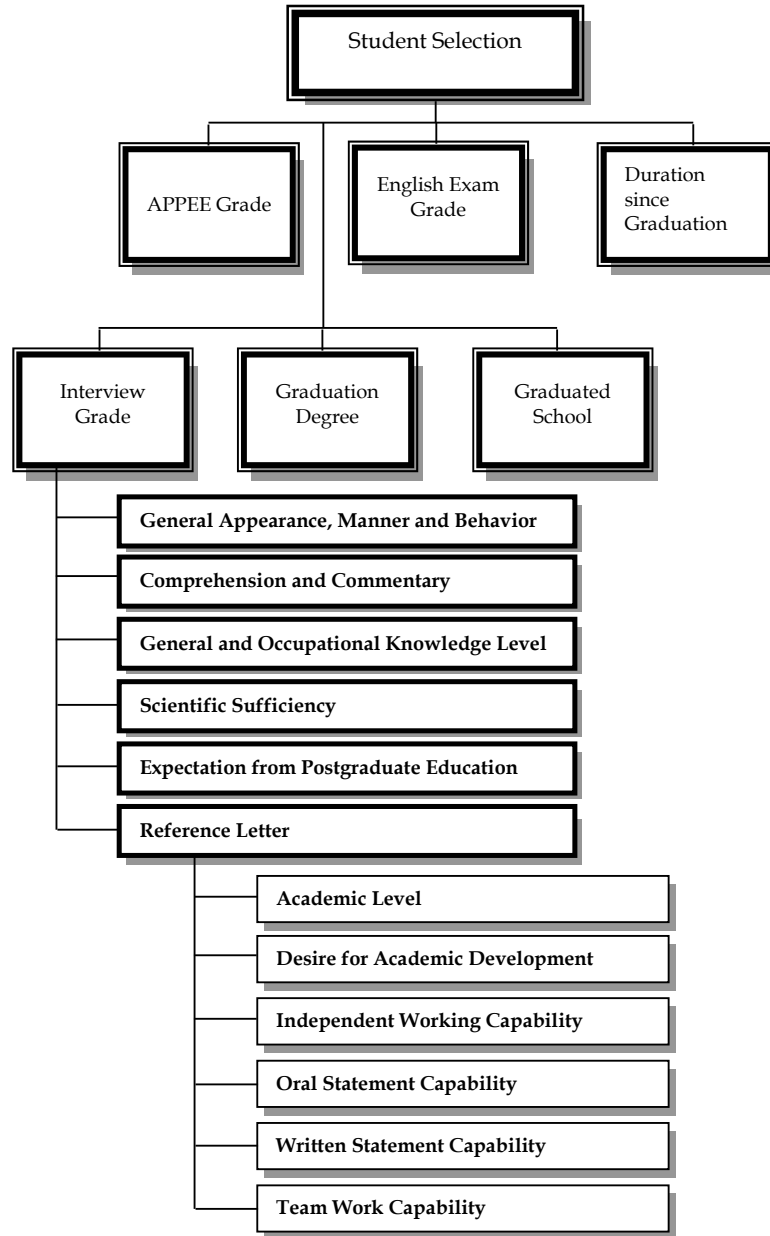


Figure 1. Hierarchical Representation of Selection Criteria

Table 3*The Weights Computed By The Mathematical Model*

Main Criteria	Weight
APPEE Grade	0.500
English Exam Grade	0.176
Interview Grade	0.145
Graduation Degree	0.089
Graduation School	0.053
Duration Since Graduation	0.037

Since there are no additional constraints imposed for the subcriteria, we use the Expert Choice for the calculation of the remaining weights. The weights for the subcriteria of interview and reference letter are given in Tables 4 and 5, respectively.

Table 4*The Weights of The Subcriteria of Interview*

Subcriteria	Weight
General Appearance, Manner and Behavior Grade	0.181
Comprehension and Commentary Grade	0.246
General and Occupational Knowledge Level Grade	0.097
Scientific Sufficiency Grade	0.277
Expectation from Postgraduate Education Grade	0.120
Reference Letter Grade	0.079

Table 5*The Weights of The Subcriteria of Reference Letter*

Subcriteria	Weight
Academic Knowledge Level Grade	0.182
Desire for Academic Development	0.132
Independent Working Capability Grade	0.149
Oral Statement Capability Grade 0	0.198
Written Statement Capability Grade	0.162
Team Work Capability Grade	0.177

Duration after graduation was described as the decreasing linear function ($f(x) = 1 - 0.033x$). It is assumed that knowledge would be forgotten completely 30 years after graduation. 7 alternatives (6 universities that were found to be notable by group decision, and remainders were collected as others) were constituted from graduation school criteria's alternatives to digitize. Numerical values were obtained by pairwise comparison of these alternatives. The quintet scale is used to evaluate the reference criteria's subcriteria. Pairwise comparison is used to digitize this scale too.

In the empirical study, the scores calculated by each method and the CGPA of students who were enrolled in 2004, 2005 and 2006 are computed and ranked. The results are given in Table 6. The total deviation of scores calculated by each method from CGPA of 31 students who were enrolled at these years was computed as follows;

Step 1: "0-1 normalization" is applied to students' scores calculated by each method and the CGPA due to scale differences

Step 2: The distances between normalized scores calculated by each method and the normalized CGPA were computed and squared.

Step 3: Total deviation was computed for each method by adding the squared distance.

Table 6
Deviations Between Student Scores Calculated by Each Method and The CGPA

Student	AHP	TOPSIS	WP
S1	0.007	0.164	0.088
S2	0.004	0.014	0.074
S3	0.257	0.002	0.078
S4	0.000	0.225	0.089
S5	0.030	0.044	0.003
S6	0.585	0.037	0.134
S7	0.026	0.720	0.407
S8	0.073	0.000	0.001
S9	0.008	0.212	0.119
S10	0.002	0.427	0.191
S11	0.000	0.032	0.107
S12	0.167	0.003	0.062
S13	0.010	0.295	0.180
S14	0.010	0.338	0.130
S15	0.077	0.795	0.483
S16	0.011	0.448	0.165
S17	0.017	0.443	0.160
S18	0.782	0.782	0.782
S19	0.144	0.240	0.165
S20	0.079	0.096	0.058
S21	0.189	0.266	0.328
S22	0.007	0.002	0.002
S23	0.000	0.008	0.004
S24	0.027	0.004	0.015
S25	0.003	0.000	0.008
S26	0.122	0.175	0.134
S27	0.519	0.650	0.567
S28	0.491	0.588	0.505
S29	0.269	0.366	0.279
S30	0.921	0.697	0.906
S31	0.469	0.469	0.469
Total Deviation	5.307	8.540	6.693

Note: S_i represents students who were approved in 2004, 2005 and 2006

The total deviations for WP, TOPSIS and AHP were 6.693, 8.540 and 5.307, respectively. AHP method obtained the least deviation. Therefore candidates' scores obtained by AHP are closer to students' CGPA than other methods.

Conclusion

Student selection for postgraduate education requires the multicriteria assessment of the candidates. MCDM has shown advantages in ranking the performance of a set of decision alternatives with respect to multiple criteria in various decision contexts. In this study, the postgraduate student selection problem is described as a MCDM problem. Three MCDM methods were implemented: Weighted Product, AHP and TOPSIS. The total deviation between students' CGPA and the scores calculated by each method was computed to determine the most appropriate method. AHP

method was found to be the most appropriate method as it obtained the least total deviation. TOPSIS, which requires several computations, performed worst.

In the upcoming academic years, the Defense Sciences Institute is considering utilizing AHP for post graduate student selection. The first reason is, for sure, the empirical evidence shown in this paper. The second reason is that AHP is a simple, easy to implement, and easy to express technique.

The performance of the selected technique is tested with the data of the past three years. However, the performance of these techniques should be monitored over the years and should be replaced by better performing processes if necessary. Further research can compare some other multicriteria decision-making techniques for this problem as well. In the comparison of the techniques, several student inputs and only one student output, CGPA, is used. An enhanced comparison can be done by defining some other academic outputs of the students and performing a multiple input- multiple output analysis.

References

- Bahurmoz M.A. (2003). The analytical hierarchy process at Dar Al-Hekma, Saudi Arabia, *Interfaces*, 33, 4, July-August, 70-78
- Belton, V. and Stewart, T.J. (2002). *Multiple criteria decision analysis: An integrated approach*. Boston: Kluwer Academic Publishers.
- Blanchard, W., Pierce, G.A., and Hood, S.M. (1989). Multicriteria analysis: managing complexity a selecting a student-information system. *Research in Higher Education*, 30, (1), 21-29.
- Bridgman, P.W. (1922). *Dimensional analysis*. NewHaven, CT, Yale University Press
- Bryson, N., (1995). A goal programming method for generating priority vectors. *Journal of Operations Research Society*, (46),641-648.
- Chandran, B., Golden, B., and Wasil, E. (2005). Linear programming models for estimating weights in the analytic hierarchy process. *Computers & Operations Research*, (32), 2235-2254.
- Chu, T.C. (2002). Selecting plant location via a fuzzy TOPSIS approach. *International Journal of Advanced Manufacturing Technology* 20, 859-864
- Davey, A., Olson D., and Wallenius J. (1994). The process of multiattribute decision making: a case study of selecting applicants for a PhD program. *European Journal of Operational Research*, 72, 469-484.
- Deng, H., Yeh, C.-H., and Willis, R.J. (2000). Inter-company comparison using modified TOPSIS with objective weights. *Computers & Operations Research*, 27, (10), 963-973.
- Ertugrul I. and Karakasoglu N., (2007). Performance evaluation of Turkish cement firms with fuzzy analytic hierarchy process and TOPSIS methods. *Expert Systems with Applications*.
- Hwang, C.L. and Yoon, K. (1981). *Multiple attribute decision making: Methods and applications. A state-of-the-art survey*. New York: Springer-Verlag.

- Keeney, R.L., ve Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York, Wiley, (reprinted, New York, Cambridge University Press, 1993)
- Lai Y.J., Liu T.Y., and Hwang C.L. (1994). TOPSIS for MODM. *European Journal of Operational Research*, 76 , 486-500.
- Millet, I., (1997), Ethical decision making using the analytic hierarchy process, *Journal of Business Ethics*, 17, (11), 1197-1204.
- Mustafa, A. and Goh, M. (1996), Multi-criterion models for higher education administration. *International Journal of Management Science*, 24 (2), 167-178.
- Özpeynirci, Ö., Kazançoğlu, Y. and Altunok, T. (2009). *A linear programming based weight calculation method using pairwise comparison with side constraints*. Technical Report 0903-01, Izmir University of Economics, Faculty of Economics and Administrative Sciences.
- Parkan, C. and Wu, M.L. (1999). Measurement of the performance of an investment bank using the operational competitiveness rating procedure. *Omega, International Journal of Management Science*, 27, 201-217.
- Saaty, T.L. and Ramarujam, V. (1983). An objective approach to faculty promotion and tenure by analytical hierarchy process. *Research in Higher Education*, 18, 311-331.
- Saaty, T. L., (2001). *Decision making in complex environments: the analytic network process for decision making with dependence and feedback*, RWS Publications, USA.
- Shyur, H.J., (2006). COTS evaluation using modified TOPSIS and ANP, *Applied Mathematics and Computation* 177, 251-259.
- Starr, M.K. (1972), *Production management*. Englewood Cliffs, NJ, Prentice-Hall.
- Tanrıöğen, A., İscan, S., (2009). Time Management Skills of Pamukkale University Students and their Effects on Academic Achievement. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 35, 93-108.
- Wang Y.M. and Elhag T.M.S, (2006). Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment, *Expert Systems with Applications* 31 , 309-319.
- Wang J.J. and Yang D.L., (2007). Using a hybrid multi-criteria decision aid method for information systems outsourcing, *Computers & Operation Research* 34, 3691-3700.
- Yeh, C.-H., Deng, H. and Chang, Y.-H. (2000). Fuzzy Multicriteria analysis for performance evaluation of bus companies. *European Journal of Operational Research*, 126, (3), 459-473.
- Yeh, C.-H., (2003). The selection of multiattribute decision making methods for scholarship student selection. *International Journal of Selection and Assessment*, 11(4), December.
- Zeleny, M. (1982). *Multiple criteria decision making*. New York: McGraw-Hill.

Çok Kriterli Karar Verme Metotlarının Lisansüstü Öğrenci Seçme Problemi için Karşılaştırmalı Analizi

(Özet)

Lisansüstü öğrenim için mevcut kaynaklar, lisans için olanlarla karşılaştırıldığında oldukça sınırlıdır. Her lisansüstü öğrenci için bir akademik danışmak gerekmektedir. Akademik personel zamanının önemli bir bölümünü öğrencilerinin lisansüstü çalışmaları için ayırır. Akademik personel sayısı sınırlı olduğu için, lisansüstü öğrenci sayısı da sınırlı olmalıdır. Ayrıca bir lisans dersine birçok öğrenci katılabilir, bu, bazı lisansüstü derslerde, özellikle küçük öğrenci grupları için tasarlanmış olanlarda, mümkün olmayabilir. Bu sebeple, lisansüstü öğrencilerin seçimi önemli bir problemdir. Lisansüstü öğrenci seçimi problemi bir lisansüstü programa başvuran adaylar arasından bir öğrenci alt kümesi seçmek olarak tanımlanabilir. Doğal olarak, bu problem bir çok kriterli karar verme problemdir, çünkü her adayın birçok özelliği vardır ve karar sürecinde bu özellikler aynı anda göz önüne alınmalıdır.

Araştırmanın Amacı: Bu çalışmanın amacı alternatifleri sıralamak için geliştirilmiş farklı çok kriterli karar verme metotlarının performanslarını lisansüstü öğrenci seçme probleminde karşılaştırmaktır. Ayrıca, en iyi yöntem Kara Harp Okulu Savunma Bilimleri Enstitüsü'nün öğrenci seçimi sürecinde kullanılacaktır.

Araştırmanın Yöntemi: Çalışmada alternatifleri sıralamak için geliştirilmiş olan üç çok kriterli karar verme metodu, Analitik Hiyerarşi Süreci (AHP), Ağırlıkçı Çarpım (WP) ve İdeal Çözüme Benzerliğe Göre Sıralama Yöntemi (TOPSIS) tartışılmıştır. Bir performans ölçütü geliştirilmiş ve metotları performans değerleri Kara Harp Okulu Savunma Bilimleri Enstitüsü'nden elde edilen gerçek hayat verilerini örnek vermek amacıyla kullanarak karşılaştırılmıştır.

Araştırmanın Bulguları: Bulgulara göre, AHP diğer iki rakibi ile karşılaştırıldığında daha iyidir. AHP ile elde edilen sıralama öğrencilerin programdan mezun oldukları ağırlıklı genel not ortalamalarıyla yapılan sıralamayla oldukça benzerdir.

Araştırmanın Sonuçları ve Önerileri : Lisansüstü öğrenci seçme problemini çok kriterli karar verme problemi olarak görmek çok önemlidir ve bu çok farklı güçlü yönleri olan lisansüstü öğrencilerin programa kabul edilmesi sağlar. Bu makale farklı çok kriterli karar verme metotlarının lisansüstü öğrenci seçme problemindeki performansı tek bir performans ölçütü kullanılarak karşılaştırılmıştır. İleriki bir çalışma benzer bir karşılaştırmayı birden fazla performans ölçütü kullanarak yapmak olabilir.

Anahtar Sözcükler: Çok kriterli karar verme, lisansüstü öğrenci seçimi, AHP, TOPSIS.

The Relationship between Demographics, Self Efficacy, and Burnout among Teachers

Nilay T. Bümen*

Suggested Citation:

Bümen, N.T. (2010). The relationship between demographics, self efficacy, and burnout among teachers. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 40, 17-36.

Abstract

Problem Statement: Teacher burnout could be a problem with potentially serious consequences for the teaching careers of the teachers concerned as well as for the learning outcomes of their students. The symptoms of burnout would be reduced in environments in which teachers experience professional growth, self-efficacy, and perceived success in their career progression. In recent years, an increasing number of burnout researchers have been utilising the self-efficacy theory in their studies. Accordingly, intervention efforts should centre on providing opportunities for individual development, perhaps through the enhancement of efficacy beliefs.

Purpose of the study: This study assessed the demographics, three components of burnout (emotional exhaustion, depersonalization, and personal accomplishment), and teacher self-efficacy.

Methods: The participants were 179 primary and 622 subject teachers in Izmir, Turkey. An anonymous self-report questionnaire which is comprised of two scales (the adapted Turkish version of the Teachers' Sense of Efficacy Scale and Maslach Burnout Inventory), and demographic background information served as the research tools in this study. Data processing was carried out using SPSS-13.0 and LISREL software, and factor analysis, reliability test, frequency check, correlation coefficient, analysis of variance (ANOVA), multivariate analyses of variance (MANOVA) and multiple regressions were used while analysing the data.

Findings: Results showed that the combined measures of burnout were significantly related to teaching experience, type of school (public or private), and class size. Additionally, the perceived sense of self-efficacy was inversely correlated with perceived burnout: the lower the sense of self-efficacy, the higher the perceived burnout. The three efficacy beliefs were also the significant predictors of personal accomplishment. Efficacy for student engagement was the only variable that contributed statistically

* Associate Prof. Dr., Ege University, Faculty of Education, nilay.bumen@ege.edu.tr

significantly to emotional exhaustion. The salience of efficacy for student engagement was noted as the sole major contributor in predicting all three burnout components.

Conclusions: The findings generally supported the current research on the roles of teacher background variables and self-efficacy in burnout. It may also have practical implications for preparing teachers to avoid or reduce burnout.

Keywords: Self-efficacy, teacher self-efficacy, burnout, teacher burnout, demographics

Bandura's theory of perceived self-efficacy (1997) is a usable conceptual framework for studying the impact of emotionally charged relationships on burnout (Leiter, 1992; Cherniss, 1993). Perceived self-efficacy refers to the person's belief in his or her ability to organise and execute the courses of action required to achieve goals (Bandura, 1997). When applied to the field of education, *teacher self-efficacy* refers to teachers' beliefs in their ability to affect student outcomes. In a review of research on teacher self-efficacy, Ross (1995) concluded that teachers who believe that they are professionally effective set more challenging goals for themselves and their students, take more responsibility for student outcomes, and persist when faced with obstacles to learning.

An increasing number of researchers draw on self-efficacy theory in their research on burnout. Leiter (1992, p.43) stated that burnout is "a crisis in self-efficacy", and Cherniss (1993, p.234) wrote about "the role of professional self-efficacy in the etiology and amelioration of burnout". Many other studies demonstrate that doubts about self-efficacy can in themselves trigger the burnout process. Chwalisz, Altmaier and Russell (1992) have found that teachers who score low in self-efficacy reported a higher degree of burnout than their counterparts who score high in self-efficacy, while Greenglass and Burke (1988) have concluded that doubts about self-efficacy contributed significantly to the development of burnout among male teachers.

Despite the voluminous body of research on teacher burnout and self efficacy in Western countries, there is a scarcity of research work on these concepts in Turkey. The present cross-sectional study seeks to investigate the association between burnout, self-efficacy and demographic variables in primary and secondary school teachers from Turkey, and investigates the relative contribution of these variables on the three dimensions of burnout - emotional exhaustion, depersonalization and reduced personal accomplishment.

Teacher Self-efficacy

Teacher self-efficacy has been defined as "teachers' belief or conviction that they can influence how well students learn, even those who may be difficult or unmotivated" (Guskey & Passaro, 1994, p.4). Thus, in general, teachers' self-efficacy beliefs appear to affect the effort teachers invest in teaching, their level of aspiration, and the goals they set. Teachers with a strong sense of efficacy tend to exhibit greater

levels of planning, organisation, and enthusiasm (Allinder, 1994). Teachers with higher efficacy judgments tend to be more open to new ideas, more willing to experiment with new methods to better meet the needs of their students (Cousins & Walker, 2000). Efficacy beliefs influence teachers' persistence when things do not go smoothly and their resilience in the face of setbacks (Gibson & Dembo, 1984). A greater sense of efficacy enables teachers to be less critical of students who make errors (Ashton & Webb, 1986), to work longer with a student who is struggling (Gibson & Dembo, 1984), and makes the teachers more likely to stay in teaching (Glickman & Tamashiro, 1982). Despite the extensive study on teacher self-efficacy in Western countries, a limited number of attempts have been made to examine this important construct in non-Western contexts (e.g., Chan, 2007). Similarly, in recent years the number of studies related to teacher self-efficacy has increased in Turkey, as well (Senemoğlu, Yağcı, Demirel & Üstündağ, 2006; Ekici, 2006; Demir, 2008; Kan, 2009; Akbaş, 2010).

Teacher Burnout

Maslach and Jackson (1981) described burnout as a psychological syndrome of emotional exhaustion, depersonalization, and reduced personal accomplishment that can occur among individuals who work with other people. Emotional exhaustion refers to feelings of being emotionally overextended and depletion of one's emotional resources. Depersonalization refers to a negative, callous, and detached attitude towards the people that one works with, i.e., patients, clients, or pupils. Reduced personal accomplishment, on the other hand, is a person's negative self-evaluation in relation to his or her job performance (Schaufeli et al., 1993).

Many studies have shown that teacher burnout has serious consequences for teachers, students, schools and the entire society. For instance, Cherniss (1993) reported that teacher burnout leads to a significant decrease in the quality of teaching, long absenteeism, and early withdrawal from the profession. Teachers who become burned out may be less sympathetic toward students, have lower tolerance for frustration in the classroom, plan for their classes less often or less carefully, frequently feel emotionally and physically exhausted, anxious, irritable, and depressed, and, in general, be less committed to their work (Farber, 1984).

In earlier studies on teacher burnout, a great deal of attention has been paid to the influence of personal characteristics such as gender, age, the length of teaching experience, and the post held in the school (Hodge et al., 1994; Laughlin, 1984). Anderson and Iwanicki (1984) have found that male teachers score significantly higher in emotional exhaustion, depersonalization and personal accomplishment, yet Zhao and Bi (2003) have claimed that no significant difference exists between genders. Besides gender, age is also a salient differentiating variable for emotional exhaustion (Byrne, 1994). Marital status, age, students' ability, subjects taught, and grade level of students have also been identified as factors which relate to teacher burnout in previous studies (e.g., Anderson & Iwanicki, 1984; Hodge et al., 1994). Mo (1991) has found that the burnout phenomenon was quite evident among secondary school teachers in Hong Kong. There are some studies that were conducted about

teacher burnout in Turkey as well. For instance, while Peker (2002) states that the level of teacher burnout varies significantly according to educational background, gender, and class grade, Kırılmaz, Çelen and Sarp (2003) claim that gender, age, educational background, teaching experience, and class size do not affect teacher burnout except for job satisfaction and reason for selecting the teaching profession.

Teacher Self-efficacy and Burnout

In recent years, an increasing number of burnout researchers have been utilising the self-efficacy theory in their studies (e.g., Friedman, 2003; Chan, 2007). Professional self-efficacy was found to be positively associated with personal accomplishment, and negatively with depersonalization and emotional exhaustion, and individuals with lower levels of generalised self-efficacy have shown greater emotional exhaustion, while those with low professional self-efficacy displayed greater cynicism and less organisational commitment. Brouwers and Tomic (2000) found that self-efficacy beliefs regarding classroom management are closely related to levels of teacher burnout, in the sense that teacher self-efficacy has a longitudinal effect on depersonalization, and a synchronous effect on personal accomplishment. Evers, Brouwers and Tomic (2002) have found that self-efficacy beliefs are adversely affected by feelings of depersonalization and emotional exhaustion, and constructively by personal accomplishment.

Likewise, Chwalisz, Altmaier and Russell (1992) have found that perceived occupational inefficacy is an important cause of burnout. While teachers with high perceived efficacy direct their efforts to solving problems when faced with academic stressors, teachers who doubt their efficacy instead turn their efforts inwards to relieve their emotional distress. Therefore, a crisis in professional efficacy can be an essential contributing factor in burnout, and thus, burnout may be considered a breakdown in the occupational domain of a person's sense of his or her own efficacy (Leiter, 1992). Teachers with a low sense of efficacy are also found to be the ones most likely to drop out of the teaching profession (Glickman & Tamashiro, 1982), and results reveal that teachers with high levels of self-efficacy and more resources suffer less stress and burnout than teachers with low self-efficacy and fewer resources, and vice versa (Chan, 2002; Betoret, 2006), which suggests that when teachers have a positive perception of their self-efficacy and when resources are accessible, teacher burnout is mitigated.

As it is seen, a variety of research has been conducted on the relationship between teacher self-efficacy and burnout, yet, the role of self-efficacy for burnout is difficult to catch, as reduced personal accomplishment (i.e., self-efficacy) is part of the burnout definition. On the other hand, changes in self-efficacy might be foretokens of burnout development. Hence, it is sensible to think that more research needs to be done on the question of whether self-efficacy has a role as a mediator or moderator in a stress-burnout relationship.

In recent years, the Turkish educational system has become the target of widespread scrutiny and criticism, propelling incessant waves of education reform. Having to implement these reform measures, teachers are pressed to do more work

with no additional or even fewer resources, while receiving fewer rewards and less recognition for their efforts. It is therefore no surprise that teacher burnout in Turkey has become a popular topic for research studies (e.g., Sari, 2004; Deryakulu, 2005). Furthermore, the implementation of new Turkish primary school curriculum reform threatens the teacher self-efficacy in several ways. For example, the reform asks that teachers (a) implement unfamiliar instructional strategies, (b) draw on alternative assessment knowledge they may not have, and (c) use new instructional materials (Gözütok, Akgün & Karacaoğlu, 2005; Bukova-Güzel & Alkan, 2005). Teacher burnout could be a problem with potentially serious consequences for the teaching careers of the teachers concerned as well as for the learning outcomes of their students. Thus, the considerable implications for teachers' well-being, and their performance relative to students and colleagues, cannot be overlooked. In this connection, Kelchtermans and Strittmatter (1999), for example, have suggested that the symptoms of burnout would be reduced in environments in which teachers experience professional growth, self-efficacy, and perceived success in their career progression. Accordingly, intervention efforts should centre on providing opportunities for individual development, perhaps through the enhancement of efficacy beliefs.

This is the first study examining the relationship between demographic variables, self-efficacy and burnout in a sample of in-service teachers in Turkey. Specifically, three questions have been raised in an attempt to find out the self-efficacy and burnout situation among teachers: (1) Is there a relationship between teachers' demographic variables (e.g., gender, teaching experience, class size, weekly course load, etc.) and burnout? (2) Is there a relationship between teachers' self-efficacy and burnout? (3) What is the best predictor among the self-efficacy factors for each of the burnout components? By answering these questions, this study can expand our understanding of the teacher self-efficacy and burnout relationship beyond Western contexts by providing information about Turkish teachers and can make considerable contributions to teacher education in Turkey. In addition, since it seeks to uncover the relations with demographic variables, this study can help identify those who are at higher risk and devise appropriate means to help them cope better with the situation. Furthermore, because of the limited amount of research on teacher self-efficacy and burnout in Turkey, this study attempts to fulfil this gap by extending teacher burnout studies by investigating its relationship with self-efficacy.

Method

Sample

There are 1145 (primary and secondary) schools and a total of 26623 teachers in Izmir, Turkey. 30 primary and secondary school teachers were invited to participate in this study, and were randomly selected according to the type of school where they worked (i.e., eight private schools and 22 public schools). During its execution, 897 out of the 1000 total questionnaire printouts were returned, which stood for a 98% response rate, and 96 were identified with organised or missing answers. Thus, after the

elimination of defective responses, 801 teachers constituted the sample of this study. Of the participants, 250 were male (31.5%), 551 (68.5%) were female, and 179 were primary teachers while 622 were subject teachers. The mean number of years of teaching experience, which ranged from 1 to 39, was 18.14 ($SD = 8.86$), and the mean number of weekly course load (as hours) was 23.18 ($SD = 5.71$). More than three quarters of the participants had finished professional training programs (77.91%). The mean number of class size of the schools was 29.41 ($SD = 9.41$). Regarding the demographic characteristics, this teacher sample was heterogeneous in nature, and was similar to the teacher population of recent years (Turkish National Education Department, 2008).

Instrumentation

An anonymous self-report questionnaire, which was comprised of two scales, and demographic background information served as the research tools in this study. Background information was requested on gender, years of teaching experience, type of school, class grade, class size, weekly course load, level of education, and academic degree.

Turkish Teachers' Sense of Efficacy Scale (TTSES). The original TSES was developed by Tschannen-Moran and Woolfolk Hoy (2001), which was based on the scale of Bandura (2001). The TSES contains 24 items comprising three subscales (each of eight items): *Efficacy for student engagement*, *efficacy for instructional strategies*, and *efficacy for classroom management*. The adapted Turkish version of the TSES was developed by Çapa, Çakıroğlu and Sarıkaya (2005), by using a 9-point scale ranging from 1-nothing, 3-very little, 5-some influence, 7-quite a bit, and 9-a great deal. However, Çapa, Çakıroğlu and Sarıkaya (2005) conducted the confirmatory factor analysis (CFA) on the *preservice* teachers; then CFA was employed to model a three factor solution using LISREL Version 8.70 (Jöreskog & Sörbom, 1993) for the 801 teachers in this study (see Path Diagram I in Appendix 1). Goodness of fit indexes ($\chi^2 = 775.74$, NNFI = 0.98, AGFI = 0.90, GFI = 0.93, RMSEA = 0.05) higher than 0.90 indicate a good fit and the RMSEA = .05 indicates a moderate fit of the model (Schumacher & Lomax, 2004; Jöreskog & Sörbom, 1993). Reliability estimates for subscale scores were $\alpha = 0.87$ for student engagement; $\alpha = 0.78$ for instructional strategies; $\alpha = 0.89$ for classroom management; and $\alpha = 0.93$ for the whole scale score.

Teacher Burnout Scale. An adaptation of the Maslach Burnout Inventory (MBI) (Maslach & Jackson, 1981) contains 22 items comprising three subscales: emotional exhaustion, personal accomplishment, and depersonalization, and a Turkish adaptation of it has been developed by Ergin (1992). High scores on the emotional exhaustion scales and depersonalization, and low scores on the personal accomplishment scale indicate burnout. Since the adapted Turkish version of the MBI has not been examined by factor analysis, CFA was employed (see Path Diagram 2 in Appendix 2) to model a three factor solution using LISREL Version 8.12a (Jöreskog & Sörbom, 1993). Thus, CFA based on burnout data for 801 teachers was conducted in order to model a three factor solution as suggested by Ergin (1992). According to findings, it can be said that goodness of fit indexes ($\chi^2 = 687.25$, NNFI = 0.96, AGFI = 0.91, GFI = 0.93, RMSEA = 0.05) higher than 0.90 indicate a good fit and the RMSEA of about .05 indicates a moderate fit of the model (Schumacher & Lomax, 2004; Jöreskog & Sörbom, 1993). Reliability estimates for subscale

scores were $\alpha = 0.84$ for emotional exhaustion; $\alpha = 0.82$ for personal accomplishment; $\alpha = 0.70$ for depersonalization; and $\alpha = 0.87$ for the whole score.

Procedure

The present study was conducted during the spring semester of the 2006-2007 academic year. Participants were asked to voluntarily fill out scales, and each subject was given the scales by the researcher and her volunteer graduate students directly, together with a set of specific instructions describing the study, including directions for completing the scales and returning them to the researcher. The subjects were assured of their anonymity and confidentiality concerning their responses. Following their completion, the researcher collected the scales and checked that all responses were properly scored.

Statistical Analyses

Data processing was carried out using SPSS-13.0 and LISREL software, and factor analysis, reliability test, frequency check, correlation coefficient, analysis of variance (ANOVA), multivariate analyses of variance (MANOVA) and multiple regressions were used while analysing the data.

Results

Teacher Burnout and Demographic Variables

The effects of demographic variables, including the participant's gender, type of school, teaching class grade, teaching experience, class size, weekly course load, and educational background were examined in order to study their relationship with teacher burnout in Turkey by using multivariate analysis of variance, and they were regarded as independent variables while the three burnout components were regarded as dependent variables. With the use of Wilk's criterion, the combined measures of burnout were seen to be significantly related to type of the school [$\Lambda = 0.971$, $F(3,781) = 7.822$, $p < 0.001$], class grade [$\Lambda = 0.988$, $F(3,781) = 3.287$, $p < 0.001$] and teaching experience [$\Lambda = 0.962$, $F(18,2349) = 1.670$, $p < 0.05$]. However, they were not significantly related to gender [$\Lambda = 0.991$, $F(3,78) = 2.457$, $p > 0.05$], class size [$\Lambda = 0.997$, $F(6,1564) = 0.345$, $p > 0.05$], weekly course load [$\Lambda = 0.995$, $F(6,1564) = 0.708$, $p > 0.05$], and educational background [$\Lambda = 0.978$, $F(12,2349) = 1.469$, $p > 0.05$]. Further analyses based on a univariate F -test were then carried out, such as the Tukey test, which was conducted to examine group differences for those independent variables with two or more levels. The results are shown in Table 1.

As can be seen in Table 1, type of school, class size and teaching experience differences were found in all three burnout components. Public school teachers were significantly more burned-out in all three components than their peers at private schools, and the grades 1-5 teachers tended to show greater emotional exhaustion and personal accomplishment than their grades 6-11 colleagues. Also, with respect to teachers' educational background, teachers who held B.A degrees scored higher in emotional exhaustion than other teachers. Besides, no significant difference was found in all three burnout components among teachers' gender and weekly course load.

Table 1
ANOVA of the MBI Subscales by the Demographic Variables of Turkish Teachers

Variables	N	Emotional exhaustion			Depersonalization			Personal accomplishment		
		M	SD	F	M	SD	F	M	SD	F
Gender										
Female	551	9.73	5.70	3.75	2.50	2.48	3.57	7.56	4.06	.22
Male	250	10.60	6.09		2.88	2.93		7.42	3.93	
Type of the school										
Private	250	7.85	5.30	52.47***	1.90	2.16	27.87***	6.07	3.66	50.00***
State	551	10.98	5.81		2.95	2.76		8.17	4.00	
Class grade										
Grades 1 - 5	179	8.77	5.49	10.45**	2.39	2.63	1.84	6.83	4.56	6.74*
Grades 6 - 11	622	10.36	5.89		2.69	2.63		7.72	3.82	
Teaching experience										
1- 5 years	116	8.87	6.62	3.13**	2.84	2.84	2.44*	7.20	2.95	5.49***
6- 10	138	10.23	5.46		2.90	2.67		7.84	3.70	
11- 15	172	10.68	5.97		2.70	2.56		7.89	3.64	
16- 20	126	11.15	5.54		2.81	2.59		8.59	4.82	
21- 25	99	10.10	5.88		2.45	2.86		7.29	3.95	
26- 30	92	9.22	5.34		2.40	2.55		6.90	4.30	
31 and more	58	8.29	5.29		1.51	1.85		5.32	4.45	
Class size										
1- 20 students	182	8.31	5.02	10.66***	2.03	2.22	6.20**	6.62	4.00	6.27**
21- 35	443	10.35	6.08		2.76	2.70		7.70	3.93	
36 and more	176	10.88	5.67		2.89	2.78		7.99	4.13	
Weekly course load										
1- 20 lessons	226	10.11	6.13	1.39	2.69	2.70	.99	7.67	3.89	1.50
21- 29	433	9.73	5.73		2.51	2.51		7.30	3.84	
30 and more	142	10.66	5.68		2.85	2.87		7.92	4.65	
Educational background										
Teacher school (high school)	22	9.13	6.22	3.31*	2.18	3.45	2.21	7.31	6.75	1.18
Institute of Education (associate's degree)	99	8.75	4.97		1.94	2.02		6.78	4.34	
Education College (associate's deg)	62	8.64	5.51		2.56	2.75		7.29	5.05	
Faculty of Education (B.A)	345	10.68	6.12		2.78	2.65		7.75	3.74	
Other (other faculties and conservatory)	273	9.97	5.70		2.71	2.68		7.56	3.67	

Notes: N =Number of subjects, M=Mean, SD=Standard deviation, F=F- value, *p<0.05, **p<0.01, ***p<0.001

Self Efficacy and Burnout

Pearson product-moment correlations were calculated for burnout and self-efficacy variables. An important finding in Table 2 is the generally moderate correlations between the self-efficacy and burnout variables, which range from $r = -0.24$ to $r = -0.46$. Nevertheless, all correlations are in the anticipated (i.e., negative) direction, showing that the lower the sense of self-efficacy, the higher the perceived sense of burnout (or vice versa). The correlation between emotional exhaustion and

instructional strategies ($r = -0.26$), personal accomplishment and student engagement ($r = -0.46$), and depersonalization and classroom management ($r = -0.29$) are of some salience (see Table 2).

Table 2
Means, Standard Deviations, and Intercorrelations between Burnout and Self-Efficacy Measures

Measures	M	SD	1	2	3	4	5	6	7
Burnout variables									
1. Burnout (total score)	20.15	10.00	-	0.88**	0.74**	0.73**	-0.44**	-0.38**	-0.40**
2. Emotional exhaustion	10.00	5.84		-	0.55**	0.38**	-0.29**	-0.26**	-0.27**
3. Depersonalization	2.63	2.64			-	0.38**	-0.31**	-0.24**	-0.29**
4. Personal accomplishment	7.52	4.01				-	-0.46**	-0.41**	-0.41**
Self-efficacy variables									
5. Student Engagement	54.82	7.75					-	0.74**	0.74**
6. Instructional Strategies	58.59	7.59						-	0.72**
7. Classroom Management	58.19	7.58							-

** $p < 0.01$

Standard multiple regressions were performed in order to probe for the cumulative effect of self-efficacy variables (continuous variable) on burnout variables. R for regression was significantly different from zero for burnout: $R = 0.45$ (adj. $R^2 = 0.20$) $F(3,797) = 68.64, p < 0.001$; for emotional exhaustion: $R = 0.30$ (adj. $R^2 = 0.09$), $F(3,797) = 27.17, p < 0.001$; for personal accomplishment: $R = 0.32$ (adj. $R^2 = 0.22$), $F(3,797) = 78.83, p < 0.001$ and for depersonalization: $R = 0.48$ (adj. $R^2 = 0.10$), $F(3,797) = 30.48, p < 0.001$.

Only two of the independent variables contributed significantly to the prediction of burnout, emotional exhaustion, and personal accomplishment in the regression equation, which were student engagement and classroom management. Efficacy for student engagement was the only variable that contributed statistically significantly to emotional exhaustion. All three efficacy beliefs were the significant predictor of personal accomplishment (also see Table 3). Finally, it is worth noting here that efficacy for student engagement was the sole major contributor to the statistical prediction of burnout and its components.

Table 3
*Regression Analysis Summary for Self-Efficacy Variables Predicting
 Burnout, Emotional Exhaustion, Unaccomplishment and Depersonalization (n = 801)*

Variable	R	B	SEB	β	t
Dependent variable					
Burnout (total score)	0.45				
Independent variables					
Student Engagement		- 0.36	0.06	- 0.28	- 5.44***
Instructional Strategies		- 0.09	0.06	- 0.07	- 1.38
Classroom Management		- 0.18	0.06	- 0.13	- 2.69**
Dependent variable					
Emotional Exhaustion	0.30				
Independent variables					
Student Engagement		- 0.13	0.04	- 0.18	- 3.28**
Instructional Strategies		- 0.04	0.04	- 0.05	- 0.96
Classroom Management		- 0.07	0.04	- 0.09	- 1.75
Dependent variable					
Depersonalization	0.32				
Independent variables					
Student Engagement		- 0.07	0.01	- 0.22	- 4.02***
Instructional Strategies		0.01	0.01	0.02	0.54
Classroom Management		- 0.05	0.01	- 0.14	- 2.66**
Dependent variable					
Personal accomplishment	0.48				
Independent variables					
Student Engagement		- 0.15	0.02	- 0.29	- 5.73***
Instructional Strategies		- 0.06	0.02	- 0.11	- 2.35*
Classroom Management		- 0.05	0.02	- 0.10	- 2.14*

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Discussion and Recommendations

As mentioned above, the differences in the teaching experiences were found in all three burnout components, and the findings are consistent with those of Mo (1991), Byrne (1994), Friedman (2000), Tümkaya (1996), Sarı (2004), Gündüz (2005), Lau, Yuen and Chan (2005). Teachers start their teaching career at schools with high levels of idealism and commitment, expecting to gain favourable outcomes through good performance (Farber, 1991) yet, shortly after starting teaching, they face stressful experiences both in the classroom and in the school environment in general (Friedman, 2000). Similarly, in the Turkish context, it has been determined that novice teachers suffer burnout more because of inexperience, not getting assistance from experienced colleagues, mistakes in classroom activities and problem solving (Tümkaya, 1996), and not getting enough positive feedback from their students (Gündüz, 2005).

It has been found in this study that the type of school affects the level of teacher burnout as well, in the sense that the teachers working in private schools have low burnout, and the issue which needs to be considered here is the working conditions of private and public schools. Similarly, Karaköse and Kocabaş (2006) revealed that the principal's attitude and behaviour positively affected teachers' job satisfaction and motivation at private schools while teachers seldom agreed with this opinion in public schools. On the contrary, teachers in private schools stated that they were under stress while performing their duties. Then, why do those teachers working in private schools have less burnout although they feel themselves under stress? It can be thought that private school teachers accept the work load before applying for the job or they avoid burnout in order not to lose their job. Gündüz (2005) argues that private school teachers are less burned out because of working conditions, socio-economic conditions of these schools and parents, and the relationship between teachers and administrators. Other studies about burnout in Turkey also reveal that private school teachers are less burned out and have higher job satisfaction (Minibaş, 1992; Tümkaya, 1996; Deryakulu, 2005), and Gündüz (2005) points out that private schools employ more experienced teachers, which leads to this outcome.

Crowded classrooms create a stressful environment for teachers, since they cannot get to know all students and care for them individually, and teachers reduce the academic objectives. In this sense, the present study also found that the class size leads to a difference in the level of burnout, which is consistent with some other studies (Tümkaya, 1996; Deryakulu, 2005; Şahin, 2007). However, there are some studies that did not come to such a conclusion (Kırılmaz, Çelen & Sarp, 2003; Çimen, 2007).

The findings in the literature which state that the level of burnout differs significantly according to class grade are confirmed by this study as well (Anderson & Iwanicki, 1984; Mo, 1991; Tuğrul & Çelik, 2002; Deryakulu, 2005; Çimen, 2007). Grades 6-11 teachers are more burned out than grades 1-5 teachers, which is explained through work satisfaction (Çimen, 2007) and the stress created by adolescent students (Deryakulu, 2005). Although elementary school teachers have a greater work load, they receive positive feedback from students or parents and they are responsible for the same students for five years in Turkey, which builds intense communication and positive feedback and thus, avoids burnout.

According to this study, the level of burnout does not differ with respect to teachers' gender and weekly course load. Some studies exist which identify gender as a differentiating variable with respect to burnout (e.g., Huston, 1990) as well as those which claim that gender has no role in affecting burnout (Sucuoğlu & Kuloğlu, 1996; Baykoçak, 2002; Kırılmaz et al., 2003; Zhao & Bi, 2003; Şahin, 2007). On the other hand, while the findings related to weekly course load confirm the findings of Deryakulu (2005), they conflict with Çimen's (2007).

One of the most noteworthy findings of this study is the negative relationship between self-efficacy and burnout. Although this relationship is generally moderate, it is very important since it confirms other studies (Evers, Brouwers & Tomic, 2002;

Friedman, 2003; Betoret, 2006; Egyed & Short, 2006) and is the first in Turkey. According to Friedman (2003), because of the cyclical nature of self-efficacy, there is no way of determining whether the relationship between self-efficacy and burnout is a uni-directional one, and therefore, the question of whether the level of self-efficacy is determined by burnout or if burnout is a consequence of low self-efficacy arises. The idea that burnout affects self-efficacy is not without theoretical backing. Friedman (2000) argues that burnout is a manifestation of a sense of professional failure at work. Exposure to repeated professional failures clearly affects levels of professional efficacy (Bandura, 1997), and therefore, the general causal direction may well be from self-efficacy beliefs to burnout, with a possibility of the opposite.

A relatively strong link between self-efficacy beliefs and personal accomplishment (a component of burnout) was found, which was affected by the three self-efficacy components, and this finding corroborates previous findings in the literature (Brouwers & Tomic, 2000; Brouwers, Evers & Tomic, 2001). Such a finding can be interpreted in light of self-efficacy theory, which states that enactive mastery experiences are the most influential source of efficacy information (Bandura, 1997). Teachers who, day in and day out, feel that they are not performing well may begin to doubt their own capability to execute the courses of action required to accomplish their job goals. In this view, teachers' accomplishments form the internal source from which they build their self-efficacy beliefs.

Another noteworthy finding in this study concerns the importance of both efficacy for student engagement and classroom management in predicting all three burnout components. In other words, efficacy for instructional strategies does not predict burnout. Besides, efficacy for student engagement was the sole, major contributor to the statistical prediction of burnout and its components, for which a couple of explanations can be given: First, it can be explained by the features of the efficacy for student engagement. With the items about student engagement in TSES, efficacy in the tasks of motivating uninterested students, helping students to think critically and creatively, and encouraging students to learn and to appreciate learning, is investigated. This kind of competence makes teachers tired and burned out since it provides slow feedback, takes time, and is compelling. On the other hand, it can be thought that Turkish teachers do not get stressed or do not worry much about whether they have sufficient teaching skills or not.

Second, when teachers have little confidence in their ability to maintain classroom order, student motivation and enthusiasm, they will likely give up easily in spite of the presence of the continuous meaningful learning climate, and feel themselves ineffective in their attempts to maintain effective teaching. It is reasonable to assume that there is a connection between these feelings of ineffectiveness and a decline in perceived self-efficacy. It so happens that teachers who doubt their ability to maintain powerful learning also do less to solve the burnout problem.

The present study extends past findings on the roles of self-efficacy and demographic variables in burnout in primary and secondary school teachers in Turkey. It may also have practical implications for preparing teachers to avoid or reduce burnout. The acquisition of strategies focusing on the improvement of

teachers' skills in classroom management and student engagement – especially for public school teachers – might be of significance. Therefore, this study may be of assistance for teachers through both pre-service and in-service training on classroom management and student engagement. Indeed, there is evidence to suggest that professional development, either prior to entering the teaching profession or during teaching, seems to be an important way to dispel burnout (Friedman, 2000). Furthermore, some economic remedies (e.g., encouragement, professional development opportunities) which are being carried out in private schools should be implemented in public schools as well, and the efforts to reduce class sizes must be carried on, and crowded classes should be depopulated.

Future research should explore whether general education teachers who hold more efficacious beliefs are more likely to engage in effective educational practices than their less efficacious colleagues when faced with students with academic and behavioural problems. As noted by Wheatley (2002), observations of teachers' actual teaching have been very rare in the field of efficacy. I believe that continued research examining the relationship between teacher self-efficacy, burnout, student characteristics, and teachers' instructional practices is both warranted and essential. In addition, examining Turkish teachers' self-efficacy longitudinally and collecting quantitative data about the relationship between self-efficacy and burnout could be suggested for further studies in this field.

References

- Akbaş, A. (2010). Attitudes, self-efficacy and science processing skills of teaching certificate master's program (OFMAE) students. *Eurasian Journal of Educational Research*, 39, 1-12.
- Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education*, 17, 86-95.
- Anderson, M. B., & Iwanicki, E. F. (1984). Teacher motivation and its relationship to burnout. *Educational Administration Quarterly*, 20, 109-132.
- Ashton, P.T., & Webb, R.B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.
- Bandura, A. (1997). Self-efficacy: toward a unifying theory of behaviour change. *Psychological Review*, 84, 191-215.
- Bandura, A. (2001). *Guide for constructing self-efficacy scales*. Stanford, CA: Stanford Uni.
- Baykoçak, C. (2002). *Beden eğitimi öğretmenlerinin mesleki sorunları ve tükenmişlik düzeyleri*. [Professional problems and burnout levels of physical training teacher.] Unpublished Master's Thesis, Sakarya University, Sakarya, Turkey.
- Betoret, F. D. (2006). Stressors, self-efficacy, coping resources, and burnout among secondary school teachers in Spain. *Educational Psychology*, 26(4), 519-539.

- Brouwers, A. & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, 16, 239-253.
- Brouwers, A., Evers, W. G. & Tomic, W. (2001). Self-efficacy in eliciting social support and burnout among secondary-school teachers. *Journal of Applied Social Psychology*, 31(7), 1474-1491.
- Bukova-Güzel, E. & Alkan, H. (2005). Evaluating pilot study of reconstructed Turkish elementary school curriculum. *Educational Sciences: Theory & Practice*, 5(2), 385-425.
- Byrne, B. M. (1994). Burnout: Testing for the validity, replication and invariance of causal structure across elementary, intermediate and secondary teachers. *American Educational Research Journal*, 31(3), 645-673.
- Chan, D. W. (2007) Burnout, self-efficacy, and successful intelligence among Chinese prospective and in-service school teachers in Hong Kong. *Educational Psychology*, 27(1), 33-49.
- Cherniss, C. (1993). Role of professional self-efficacy in the etiology and amelioration of burnout. In W.B. Shaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: recent developments in theory and research*. Washington, D.C.: Taylor&Francis.
- Chwalisz, K., Altmaier, E.M., & Russell, D.W. (1992). Causal attributions, self-efficacy cognitions, and coping with stress. *Journal of Social and Clinical Psychology*, 11(4), 377-400.
- Cousins, J.B., & Walker, C.A. (2000). Predictors of educators' valuing of systemic inquiry in schools. *Canadian Journal of Program Evaluation* (Special Issue), 25-53.
- Çapa, Y., Çakıroğlu, J. & Sarıkaya, H. (2005). The development and validation of a Turkish version of teachers' sense of efficacy scale. *Education and Science*, 30(137), 74- 81.
- Çimen, S. (2007). *İlköğretim öğretmenlerinin tükenmişlik yaşantıları ve yeterlik algıları*. [Primary school teachers' burnout levels and perceived self-efficacy beliefs.] Unpublished M.A thesis, Kocaeli University, Kocaeli.
- Demir, K. (2008). Transformational leadership and collective efficacy: the moderating roles of collaborative culture and teachers' self efficacy. *Eurasian Journal of Educational Research*, 33, 93-112.
- Deryakulu, D. (2005). The examination of computer teachers' levels of burnout. *Eurasian Journal of Educational Research*, 19, 35-53.
- Egyed, C. J. & Short, R. J. (2006). Teacher self-efficacy, burnout, experience and decision to refer a disruptive student. *School Psychology International*, 27(4), 462-474.
- Ekici, G. (2006). A study on vocational high school teachers' sense of self-efficacy beliefs. *Eurasian Journal of Educational Research*, 24, 87-96.
- Ergin, C. (1992). The adaptation of Maslach Burnout Inventory and burnout among nurses and doctors. (Eds. Bayraktar, R & Dağ, İ.). *VIIth National Psychology Congress*, Ankara (pp.143- 54).

- Evers, W. J., Brouwers, A., & Tomic, W. (2002). Burnout and self-efficacy: a study of teachers' beliefs when implementing an innovative educational system in the Netherlands. *British Journal of Educational Psychology*, 72(2), 227-244.
- Farber, B. A. (1984). Stress and burnout in suburban teachers. *Journal of Educational Research*. 77(6), 325-332.
- Farber, B. A. (1991). *Crisis in education*. San Francisco: Jossey Bass.
- Friedman, I. A. (2000). Burnout in teachers: shattered dreams of impeccable performance. *Journal of Clinical Psychology*, 56(5), 595-606.
- Friedman, I. A. (2003). Self- efficacy and burnout in teaching: the importance of interpersonal-relations efficacy. *Social Psychology of Education*, 6, 191-215.
- Gibson, S. & Dembo, M. H. (1984). Teacher efficacy: a construct validation. *Journal of Educational Psychology*, 76(4), 569-582.
- Glickman, C., & Tamashiro, R. (1982). A comparison of firstyear, fifth-year, and former teachers on efficacy, ego development, and problem solving. *Psychology in Schools*, 19, 558-562.
- Greenglass, E. R., & Burke, R. J. (1988). Work and family precursors of burnout in teachers: sex differences. *Sex Roles*, 18(3/4), 215-229.
- Gözütok, D. F., Akgün, Ö. E. & Karacaoğlu, Ö. C. (2005). Evaluating of reconstructed Turkish elementary school curriculum in terms of teacher competencies. *Eğitimde Yansımalar: VIII. Yeni İlköğretim Programlarını Değerlendirme Sempozyumu*. (p.17-40). Kayseri, Erciyes Üniversitesi.
- Guskey, T. R., & Passaro, P.D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, 31, 627-643.
- Gündüz, B. (2005). The prediction of the teachers' burnout according to the irrational beliefs and vocational and personal variables. *Turkish Psychological Counseling and Guidance Journal*. 3(26), 17-31.
- Hodge, G. M., Jupp, J. J. & Taylor, A. J. (1994). Work stress, distress, and burnout in music and mathematics teachers, *British Journal of Educational Psychology*, 64, 65-76.
- Hoover-Dempsey, K., Bassler, O. & Brissie, J. (1992). Explorations in parent-school relations. *Alberta Journal of Educational Research*, 85, 287-294.
- Huston, J. (1990). Teacher burnout and effectiveness: a case study. *Education*, 110(1), 70-78.
- Jöreskog, K. G. & Sörbom, D. (1993). *Lisrel 8: Structural equation modeling with the simplis command language*. Hillsdale, NJ: Lawrence Erlbaum Associate.
- Kan, A. (2009). Effect of scale response format on psychometric properties in teaching self efficacy. *Eurasian Journal of Educational Research*. 34, 215-228.
- Karaköse, T. & Kocabaş, İ. (2006). The effect of teachers' expectations on job satisfaction and motivation in private and public schools. *Journal of Theory and Practice in Education*. 2(1), 3-14.
- Kelchtermans, G., & Strittmatter, A. (1999). Beyond individual burnout: A perspective for improved schools. In R. Vandenberghe & A.M. Huberman (Eds.),

Understanding and preventing teacher burnout: A sourcebook of international research and practice (pp. 304-314). Cambridge, UK: Cambridge Press.

- Kırılmaz, A. Y., Çelen, Ü. & Sarp, N. (2003). A study on teacher burnout with a group of primary school teachers. *İlköğretim Online*. 2(1), 2-9. Available at <<http://www.ilkogretim-online.org.tr>>
- Lau, P.S.Y., Yuen, M. T. & Chan, R.M.C. (2005). Do demographic characteristics make a difference to burnout among Hong Kong secondary school teachers? *Social Indicators Research*. 71, 491-516.
- Laughlin, A. (1984). Teacher stress in an Australian setting: the role of biographical mediators, *Educational Studies*, 10, 7-22.
- Leiter, M.P. (1992). Burnout as a crisis in self-efficacy: Conceptual and practical implications. *Work and Stress*, 6(2), 107-115.
- Maslach, C. & Jackson, S.E. (1981). Measurement of experienced burnout. *Journal of Occupational Behavior*. 2, 1-15.
- Minibaş, J. (1992). Private and state school teachers' job satisfaction levels. Paper presented at the VIIth National Psychology Congress, Hacettepe University, Ankara.
- Mo, K.W. (1991). Teacher burnout: relations with stress, personality, and social support. *Education Journal*. 19, 3-12.
- Peker, R. (2002). Factors related with burnout levels of primary school teachers. *Uludağ University Education Faculty Journal*, 15(1), 305-318.
- Ross, J.A. (1995). Strategies for enhancing teachers' beliefs in their effectiveness: research on a school improvement hypothesis. *Teachers College Record*, 97(2), 227-251.
- Sarı, H. (2004). An analysis of burnout and job satisfaction among Turkish special school headteachers and teachers, and the factors effecting their burnout and job satisfaction. *Educational Studies*, 30(3), 291-306.
- Schaufeli, W. B., Maslach, C. & Marek, T. (1993). *Professional burnout: Recent developments in theory and research*. Washington, DC: Taylor & Francis.
- Schumacher, R. E & Lomax, R.G. (2004). *A beginner's guide to structural equation modeling*. New Jersey: Lawrence Erlbaum Associates.
- Senemoğlu, N., Yağcı, E., Demirel, M. & Üstündağ, T. (2006, September). Primary school teachers' self-efficacy levels. Paper presented at the 15th National Congress on Educational Sciences, Muğla University, Turkey.
- Sucuoğlu, B. & Kuloğlu, N. (1996). Burnout levels of teachers' dealing with handicapped children. *Turkish Journal of Psychology*. 10(36), 44-60.
- Şahin, D. (2007). *Öğretmenlerin mesleki tükenmişlik düzeyleri*. [Burnout levels of the teachers]. Unpublished M.A thesis, Gazi University, Ankara, Turkey.
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783-805.
- Tuğrul, B. & Çelik, E. (2002). Primary school women teachers' burnout levels. *Pamukkale University Education Faculty Journal*. 12(2), 1-11.

- Turkish National Education Department: 2008. Teacher Statistics 2007. Turkey.
- Tümekaya, S. (1996). *Tükenmişlik görülen öğretmenlerdeki psikolojik belirtiler ve başa çıkma davranışları*. [Teacher burnout and coping strategies]. Unpublished PhD. dissertation. Çukurova University, Adana, Turkey.
- Wheatley, K. F. (2002). The potential benefits of teacher efficacy doubts for educational reform. *Teaching and Teacher Education*, 18, 5-22.
- Zhao, Y. & Bi, C. (2003). Job burnout and the factors related to it among middle school teachers, *Psychological Development and Education*, 1, 80-84.

Demografik Değişkenler, Öğretmenlerin Özyeterlik ve Tükenmişlik Düzeyleri Arasındaki İlişkiler

(Özet)

Problem Durumu: Öğretmenlerde görülen tükenmişlik, öğrencilerin öğrenme ürünlerine olduğu kadar, öğretmenlerin mesleki yaşamlarında da ciddi sonuçlara yol açabilen bir problemdir. Tükenmişliğin belirtileri öğretmenlerin mesleki gelişim, özyeterlik ve kariyer geliştirme sürecinde yaşadığı başarılarla dolu bir ortamla azalabilecektir. Son yıllarda tükenmişlikle ilgilenen araştırmacılar özyeterlik kuramından artan oranda yararlanmaktadır. Buna göre bireysel gelişim fırsatları sağlamak ya da özyeterlik inançlarının güçlendirilmesi tükenmişlikle mücadelede etkin bir rol üstlenebilir. Batı ülkelerinde öğretmenlerde görülen tükenmişlik ile özyeterlik inançları arasındaki ilişkilere dair çok sayıda araştırma bulunmasına rağmen, Doğu Avrupa ve/veya Orta Doğu'da bu konulara ilişkin pek az çalışma vardır.

Araştırmanın Amacı: Bu çalışmada demografik değişkenler, tükenmişliğin üç alt boyutu (duygusal tükenme, duyarsızlaşma ve kişisel başarı) ve öğretmenlerin özyeterlik inançları arasındaki ilişkiler incelenmiştir. Belirgin olarak üç soruya yanıt aranmıştır: 1) Öğretmenlerin cinsiyet, yaş, kıdem, sınıftaki öğrenci sayısı, haftalık ders yükü vb. özellikleriyle tükenmişlik durumları arasında anlamlı bir fark var mıdır? 2) Öğretmenlerin tükenmişlikleri ile özyeterlik inançları arasında anlamlı bir ilişki var mıdır? 3) Tükenmişliğin alt boyutlarını yordayan özyeterlik inançları hangileridir? Bu soruların yanıtlanmasıyla tükenmişlik ve özyeterlik inançları arasındaki ilişki Batı ülkeleri bağlamından genişletilerek, Türk öğretmenleri hakkında bilgi edinilebilmesi mümkün olabilecek ve öğretmen eğitimine önemli katkılar sunulabilecektir. Ayrıca demografik değişkenler açısından yapılan incelemeler, tükenmişlikte yüksek risk taşıyan grupların tanınması ve bu durumlarla uygun şekilde baş etmeyle ilgili planların yapılmasına yardımcı olabilecektir. Bunun yanında Türkiye'de tükenmişlik ve özyeterlik inancını birlikte inceleyen

çalışmaların çok az yapılması nedeniyle, araştırmanın bu alandaki boşluğun giderilmesine katkı getireceği düşünülmektedir.

Araştırmanın Yöntemi: Çalışmaya İzmir’de görev yapan 179 sınıf ve 622 branş öğretmeni (toplam 801) katılmıştır. Katılımcıların 250’si bay (%31.5), 551’i bayan (%68.5) dir. Kıdem ortalaması 18.14 (S=8.86), haftalık ders yükü ortalaması ise 23.18 (S=5.71) dir. Katılımcıların dörtte üçü (%77.91) mesleki formasyon programlarına katılmışlardır. Ortalama sınıf mevcudu 29.41 (S=9.41) dir. Çalışmada veri toplama aracı olarak iki ölçek (Öğretmen özyeterlik ölçeği -TSES & Maslach Tükenmişlik Envanteri -MBI) ve demografik değişkenleri ölçen kişisel bilgi formu kullanılmıştır. TSES’te her biri sekiz maddeden oluşan üç alt boyut ölçülmektedir: öğrenciyle kaynaşmada yeterlik, öğretim stratejilerinde yeterlik, sınıf yönetiminde yeterlik. Doğrulayıcı faktör analizi sonuçlarına göre uyum indeksleri 0.90’ın üzerindedir ($\chi^2 = 775.74$, NNFI = 0.98, AGFI= 0.90, GFI =0.93, RMSEA=0.05) ve RMSEA = .05 değeri de orta derecede uyumlu bir model elde edildiğini göstermektedir. Öğrenciyle kaynaşma alt boyutu için $a = 0.87$, öğretim stratejileri için $a = 0.78$, sınıf yönetimi için $a = 0.89$ ve ölçeğin toplam puanları için $a = 0.93$ güvenilirlik katsayıları elde edilmiştir. Maslach Tükenmişlik Envanteri ise 22 maddeden ve duygusal tükenme, duyarsızlaşma ve kişisel başarı olarak tanımlanan üç alt boyuttan oluşmaktadır. Doğrulayıcı faktör analizi sonuçlarına göre uyum indeksleri 0.90’ın üzerindedir ($\chi^2 = 687.25$, NNFI = 0.96, AGFI= 0.91, GFI =0.93, RMSEA=0.05) ve RMSEA = .05 değeri de orta derecede uyumlu bir model elde edildiğini göstermektedir. Duygusal tükenme alt boyutu için $a = 0.84$, kişisel başarı için $a = 0.82$, duyarsızlaşma için $a = 0.70$ ve ölçeğin tümü için $a = 0.87$ güvenilirlik katsayıları elde edilmiştir. Veri analizinde SPSS-13.0 ve LISREL yazılımları kullanılmış, faktör analizi, güvenilirlik katsayısı, frekans dağılımı, korelasyon, varyans analizi ve çoklu regresyon yapılmıştır.

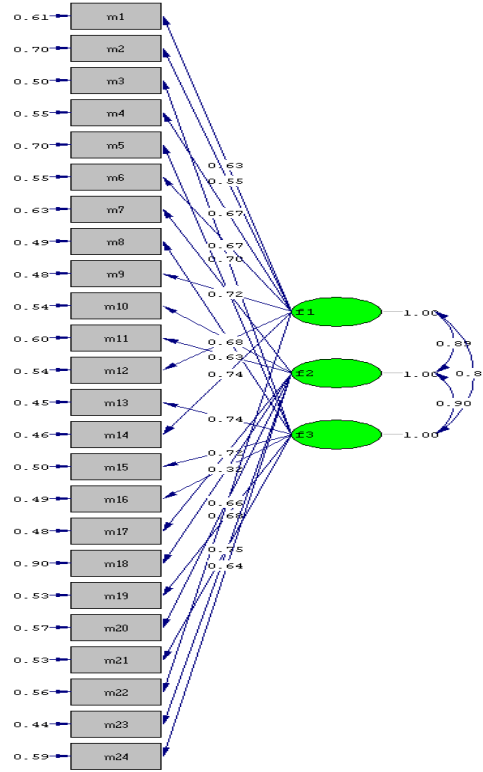
Araştırmanın Bulguları: Bulgulara göre, tükenmişliğin üç alt boyutunda da kıdem, okul türü (özel ya da devlet) ve sınıf mevcudu ile anlamlı farklılıklar elde edilmiştir. Devlet okulunda çalışan öğretmenler özel okulda çalışanlardan daha fazla tükenmektedir. Sınıf öğretmenleri duygusal tükenme ve kişisel başarı boyutlarında branş öğretmenlerine göre daha fazla tükenmektedir. Cinsiyet açısından anlamlı bir farklılık elde edilmemiştir. Özyeterlik inançları tükenmişlik durumu ile ters yönde (olumsuz) ve orta düzeyde ilişkilidir; düşük özyeterlik inancı yüksek tükenmişlik durumuna işaret etmektedir. Özyeterlik inancının üç alt boyutu da tükenmişliğin kişisel başarı boyutunun anlamlı yordayıcısıdır. Öğrenciyle kaynaşmadaki özyeterlik inancı duygusal tükenmenin tek yordayıcısıdır. Tükenmişliğin tüm alt boyutlarının (ortak) tek yordayıcısı öğrenciyle kaynaşmadaki özyeterlik inancıdır.

Araştırmanın Sonuç ve Önerileri: Çalışmanın bulguları demografik değişkenler ve özyeterlik inancının tükenmişlik üzerindeki etkilerini açıklayan diğer araştırmaları desteklemektedir. Bulguların, öğretmenlerin tükenmişlik yaşamalarını önlemek ya da azaltmak adına yapılacak girişimlere yol göstereceği düşünülebilir. Özellikle devlet okullarında

görev yapan öğretmenlerin sınıf yönetimi ve öğrenciyle kaynaşmayla ilişkili becerilerinin hem hizmet öncesi hem de hizmet içi eğitimlerle geliştirilmesine yönelik çalışmaların yararlı olacağı söylenebilir. Ayrıca özel okullarda yapılan bazı iyileştirmelerin (teşvik, mesleki gelişim olanakları vb.) devlet okullarında da uygulanması, sınıf mevcutlarının düşürülmesine yönelik çabaların sürdürülmesi de önerilebilir. Başka araştırmalarda sınıf uygulamalarına ilişkin bilgilere dayanmanın, uzunlamasına çalışmanın ve/veya nitel yöntemler kullanmanın, tükenmişlik ve özyeterlik inancı ilişkisine dair daha doyurucu bilgiler sunacağı düşünülebilir.

Anahtar Sözcükler: Özyeterlik, öğretmen özyeterliği, tükenmişlik, öğretmen tükenmişliği, demografik değişkenler

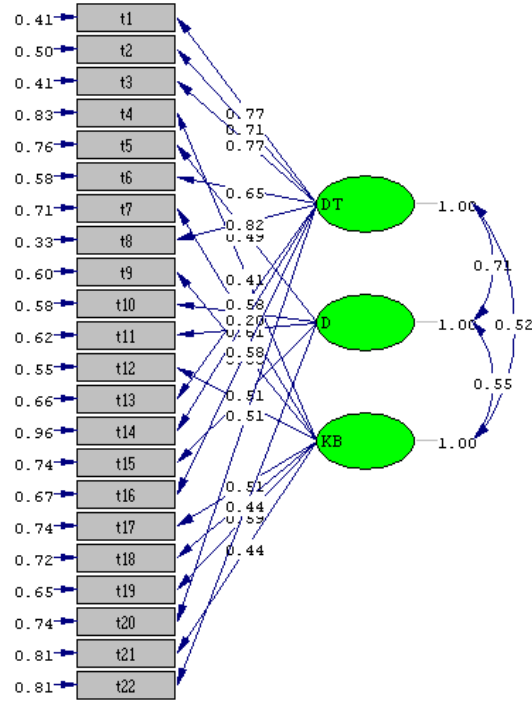
Appendix 1.



quare=775.74, df=224, P-value=0.00000, RMSEA=0.055

Figure 1. Three factor CFA model of Turkish Teachers' Self Efficacy Scale
(Note: f_1 : Efficacy for student engagement, f_2 : Efficacy for instructional strategies, f_3 : Efficacy for classroom management)

Appendix 2.



Chi-Square=687.25, df=202, P-value=0.00000, RMSEA=0.055

Figure 2. Three factor CFA model of Turkish version of the MBI
 (Note: DT: Emotional exhaustion, D: Depersonalization, KB: personal accomplishment)

Cooperative Learning Approach to Teaching Science*

Cavide Demirci*

Suggested Citation:

Demirci, C. (2010). Cooperative learning approach to teaching science. *Eğitim Araştırmaları - Eurasian Journal of Educational Research*, 40, 37-52.

Abstract

Problem Statement: In a teaching and learning environment, pupils' active participation with lessons as well as the use of learning approaches that help students remember what they learn are important dimensions. Effective learning approaches increase the quality of learning. Pupils learn best by doing and experiencing. To be effective and efficient, science teaching programs need to be pupil-centered and based on research and investigation, and should include experimental activities. In a cooperative learning approach, pupils and teachers are in a state of dynamic interaction in the classroom. When pupils interact in cooperative groups they learn to give and receive information, develop new understandings and perspectives, and communicate in a socially acceptable manner. This study investigates how a cooperative learning approach affects 7th grade primary school students' achievements and attitudes.

Purpose of Study: The aim of this research is to discover whether there is a significant difference between the means of achievement and attitude scores under a cooperative learning approach and a conventional learning approach.

Methods: An experimental method was used in the study. The "pre-test post test design with control group" method was employed.

Findings and Results: There is a significant difference in the mean value of the difference between pre- and post-test grades for a science lesson taught under a cooperative learning approach when compared to the same lesson taught using a conventional learning approach. The cooperative learning approach is more efficient than the conventional approach. The pupils taught in their science class using a conventional learning approach developed a positive attitude about the science lesson. As a result, the applications of cooperative learning were affected in both achievement and attitude.

Suggestions: Sample training circumstances for different lessons can be prepared that are appropriate for the cooperative learning approach. When evaluating the learning products, the special importance must be given to

*Assist. Prof. Dr., Eskişehir Osmangazi University, Faculty of Education, cdemirci@ogu.edu.tr.

the curricula for which the cooperative learning approach is administered. To do this, pupil's participation in class activities and portfolios can be emphasized. Descriptive studies may be carried out to determine to what extent teachers have knowledge about cooperative learning and their willingness to use cooperative learning. Descriptive studies may also be conducted on the differences in the use of cooperative learning in specific countries and abroad.

Keywords: Cooperative learning approach, science, achievement, conventional learning approach, attitude.

Cooperative learning is an old concept. Teachers have engaged their pupils in group projects, group activities, and group discussions and encouraged them for a long time. However, it is wrong to call all group-based learning cooperative learning. Cooperative learning is a successful learning approach in which small teams, each with pupils of different levels of ability, use a variety of learning activities to improve their understanding of a subject. Each member of a team is responsible not only for learning what is taught but also for helping teammates learn, thus creating an atmosphere of achievement. Pupils work through the assignment until all group members successfully understand and complete it (McLeish, 2009).

In a cooperative learning approach, pupils and teachers are in a state of dynamic interaction in the classroom (Mahran, 2000). When pupils interact in cooperative groups, they learn to give and receive information, develop new understandings and perspectives, and communicate in a socially acceptable manner. It is through interacting with each other in reciprocal dialogues that pupils learn to use language differently to explain new experiences and new realities and, in so doing, construct new ways of thinking and feeling (Gillies, 2004). Cooperative learning promotes pupils working together, achieving academic success, and developing social skills. Cooperative learning processes lessen individual competitiveness and foster cooperative small problem-solving group behavior (Johnson, Johnson & Holubec, 1998; Millis, 1995; Slavin, 1991). By having learners treat each other as resources and requiring them to go beyond a superficial engagement with learning materials, cooperative learning provides the social context for pupils to actively learn and make deeper connections among facts, concepts, and ideas (Lin, 2006). It is only under the following conditions that cooperative efforts may be expected to be more productive than competitive and individualistic efforts:

1. *Clearly perceived positive interdependence.* Jacobs (2006) defines positive interdependence as a perception among group members that what helps one group member helps all, and what hurts one group member hurts all. Positive interdependence encourages cooperation and a feeling of support. Team members perceive that they need each other in order to complete the group's task—that they "sink or swim together"; this requires each pupil in a small group to contribute to the learning of the group (Johnson, Johnson, & Smith 1998; Jolliffe 2007).

2. *Clearly perceived individual accountability and personal responsibility to achieve the group's goals.* Ghaith (2002) indicates that individual accountability exists when each individual member feels responsible to learn, to demonstrate that learning, and to

contribute to the learning of teammates. Jolliffe (2007) uses the phrase “no hitchhiking” to indicate that each member of the group is accountable for completing his or her part of the work. It is important that no one hitchhikes on the work of others. This concept requires each pupil in the group to develop a sense of personal responsibility for his or her own learning and to help the rest of the group to learn. According to Stahl (1994), teachers put pupils in cooperative learning groups so that all pupils can achieve higher academic success individually than if they were to study alone. Consequently, each pupil must be held individually responsible and accountable for doing his or her own share of the work and for learning what has been targeted to be learned (McLeish, 2009).

3. *Considerable promotive (face-to-face) interaction.* Pupils promote each others’ learning by helping, sharing, and encouraging efforts to learn. Pupils explain, discuss, and teach what they know to classmates (Ghaith, 2002; Johnson et al., 1998). According to Johnson and Johnson (1990), positive interdependence results in promotive interaction, which may be defined as individuals encouraging and facilitating each others’ efforts to achieve, complete tasks, and produce in order to reach the group’s goals.

4. *Frequent use of relevant interpersonal and small-group skills.* Webb (2002) defines social skills as group members knowing how to provide effective leadership, decision-making, trust-building, communication, and conflict management, and being motivated to use the prerequisite skills. For the cooperative learning environment to be successful teachers should teach these skills as purposefully and precisely as academic skills and the learner should utilise the skills they have learnt in completing assigned activities (Stahl 1994; Johnson et al., 1998). As Stahl (1994) points out, merely placing pupils in groups and expecting them to use appropriate social and group skills does not mean pupils will automatically use these skills. To work together as a group, pupils need to engage in such interactive abilities as leadership, trust-building, conflict management, constructive criticism, encouragement, compromise, negotiation, and clarifying. Teachers may need to describe the expected social interaction behaviors and attitudes to pupils and assign specific roles to particular pupils to ensure that they consciously work on these behaviors in their groups.

5. *Frequent and regular group processing of current functioning to improve the group’s future effectiveness.* According to Johnson et al. (1998) and Webb (2002), groups need specific time to discuss how well they are achieving their goals and whether they are maintaining effective working relationships among members. Teachers need to ensure that there is some structure to the group processing. This can be done by assigning such tasks as listing at least three member actions that help the group be successful and listing one action that could be added to make the group even more successful tomorrow. It is very important for teachers to monitor the groups and give feedback to the groups and the class as a whole on how well the groups are working together.

All healthy cooperative relationships have the above five basic elements. Researchers have focused on some cooperative learning subjects, examining and practicing them (Ding, Piccolo & Kulm, 2007). They include the following:

Importance of peer relationships. Formative peer relationships contribute to student cognitive development and socialization in many ways. Most important for this study, they help students acquire social attitudes, values and skills; develop social values and commitments; view problems from a more complex perspective; increase achievement and productivity, and influence educational and career goals.

Interaction patterns. Pupils encourage each other to achieve as a result of "primitive interaction," which is defined by personal academic approval, information sharing, and self-motivation. A teacher involves the whole class in a cooperative learning classroom, thus improving the thinking skills of the whole class, the group and the pupils. Since this improvement chain indicates that learners teach themselves, teacher interventions are a necessity, but the interactions need to be balanced in a sensitive manner.

Learning outcomes. Group efforts are more productive under the following conditions: distinct positive interdependence; frequent face-to-face interaction for assistance, information, and constructive dialogue; individual accountability; constant usage of interpersonal and small group skills; and frequent assessment of the group's effectiveness.

Relationships among outcomes. The outcome categories are interconnected and influence each other. Lasting friendships are built upon the success of joint projects. Good friends can influence each other to work hard to attain a mutual goal (Johnson & Johnson, 1990).

Cooperative learning is grounded on various theories of psychology. The idea is that man is a social being, and as such, various forms of social interaction are essential for human societal survival. Within the classroom the concept of cooperation can be promoted since individuals will be learning to work together for the overall benefit of the group (McLeish, 2009). Studies indicate that cooperative learning is useful and beneficial.

The statement, "What the child is able to do in collaboration today, he will be able to do independently tomorrow" (Vigotsky, 1987) is of significant value in establishing a rationale for cooperative learning. The aim of this research is to reveal whether there is a significant difference between the means of achievement and attitude scores under a cooperative learning approach and under a conventional learning approach. Within the guidelines of this general goal, this study sought to answer the following research questions:

1) Is there any significant difference between the mean of achievement scores under a cooperative learning approach and the mean of achievement scores under a conventional training approach in the teaching of science?

2) Is there any significant difference between the mean of post-attitude scores under a cooperative learning approach and the mean of post-attitude scores under a conventional training approach in the teaching of science?

Method

Research Design

For this research, an experimental design that provides quantitative data about the primary and secondary subproblems (called "Pre-test, Post-Test with Control Group Design") was implemented. The research design is given in Table 1.

Table 1

Pre-Test, Post-Test with Control Group Design

Groups	Pre-test	Experimental treatment	Post-test
G1	T1-T2	Cooperative learning approach in science training	T1-T2
G2	T1-T2	Conventional training approach in science training	T1-T2

G1: The group with which the cooperative learning approach was used. After pre-testing, lessons were begun. Pupils in this group were encouraged to become more active and responsible in the learning process, and used cooperative learning techniques. During learning activities, it was expected that pupils complete research, make decisions, and cooperate. The teacher encouraged the pupils by talking to them and giving rewards that caused them to ask questions of each other and do research. The teacher guided the pupils when necessary. The teacher created a highly structured, well-organised learning environment in the classroom, setting goals, planning and structuring tasks, establishing the physical arrangement of the classroom, assigning pupils to groups and roles, and selecting materials and time.

G2: The control group with which the conventional training approach was used. In the conventional approach, the teacher was active and used explanation methods. Pupils listened to the teacher. The teacher also asked questions during the process. During the teaching of science, only a blackboard was used. Lessons were taught by solving examples using the blackboard; the pupils were inactive in this approach.

T1: Achievement test about the structure of the atom prepared and developed by the researcher. The test was implemented as both a pre-test and a post-test which included 25 items. Each correct answer was given 1 point and each wrong answer was given 0 points. In the scope of this study, the aim of the lesson on the structure of the atom was determined to be the same as the aim of the lesson "Tebliğler Dergisi" from the Ministry of National Education.

The achievement test for this study was prepared with at 38 test items, with every item having four options. The test items were sent to experts to get their opinions, who made suggestions about the basis of each question, the choices, the content validity and whether the questions measured the success of pupils. After getting

these recommendations, the test items were changed and 34 questions were created. The test was given in four seventh grade classrooms having total 96 pupils. Item analysis was made at the end of this application. The item difficulty index (pj) and item separation power index (rb) for each item were calculated. A power of high separation was considered in the choice of items, and items with smaller than 0.21 points in the separation power index were eliminated. Items were corrected and included in the test if they had 0.20-0.30 points in the separation power index. If an item had more than 0.30 points they were included directly in the test. Thus, the final test form included 25 questions. Test items related to understanding, application and analysis level. KR-20 reliability of the final test was found to be 0.82.

T2: Attitude scale for the science lesson. This measure was used as both pre- and post-attitude. It was developed by the researcher using a Likert-type attitude scale with 32 items. The attitude scale was given to both the experiment and the control group before and after the experimental teaching unit.

To develop the attitude scale, a review of literature was conducted and attitude scales developed for other lessons were investigated. Later, an instrument with a 66 item Likert-type scaling was developed. Half of the questions were positively stated and half were negatively stated. The responses varied from "Completely Agree" to "Completely Disagree." For face validity, the instrument was submitted to experts in the educational sciences. The instrument was revised based on their suggestions and a 60-item scale was prepared. The instrument was used with a sample of 366 pupils from 4th through 8th grades. Following the administration of the instrument, negatively stated items were recoded. An exploratory factor analysis was conducted; only one factor emerged and the items were all higher than .40. Cronbach Alpha reliability of the instrument was .97. The highest- and lowest-scoring 27% were determined for the 366 pupil sample. T-tests were conducted for each item between the high and low 27% scorers. Finally, 32 items were selected (16 positively stated and 16 negatively stated) to become the Science attitude scale. Validity and reliability studies were conducted. The reliability of the scale was found to be .96; it was thought that the coefficient was enough for an attitude scale.

Cooperative Learning Techniques in Lesson

STAD (Student Team-Achievement Division). Pupils are assigned to four-member learning teams that are mixed in performance level, gender, and ethnicity. With STAD, the teacher presents the lesson and the pupils work within their teams to – make sure that all members have mastered the lesson. Pupils take individual quizzes on the material, at which time they may not help one another. The pupils' test scores are compared to their own past averages, and points are awarded on the basis of the degree to which pupils meet or exceed their own earlier performance. These individual points are then summed to form a team score, and teams that meet certain criteria may earn certificates or other types of rewards.

CIRC (Cooperative Integrated Reading and Composition). The most celebrated of Slavin's cooperative learning arrangement is the CIRC program for elementary and middle grade children. In CIRC groups, pupils work in four-member cooperative learning teams. They engage in a series of activities with one another, including

reading to one another, making predictions about how narrative stories will come out, summarizing stories to one another, writing responses to stories, and processing spelling, decoding, and vocabulary. They also work together to master main ideas and comprehension skills.

Jigsaw. This teaching tool is most appropriate in such subjects as social studies, literature, and related areas in which concepts rather than skills are the learning goals. The instructional raw material for Jigsaw should be a chapter, story, biography, or descriptive material. In Jigsaw, pupils work in heterogeneous teams. The teacher divides the material into sections—one for each pupil. Pupils prepare their own section of material by reading, conducting experiments, and solving problems with the help of manipulatives. The pupil's preparation might be done alone—in class or for homework—or with "preparation partners." This decision is made by the teacher, depending on the nature of the assignment and the abilities of the pupils.

Pupils meet in "practice pairs" where each pupil meets with someone from a different group who has learned the same material. The purpose of this is to both review and reconceptualise the material and to plan how the material might be best taught or presented to teammates. Pupils then present their work to other members of their groups. The teacher encourages pupils to ask questions and engage in genuine discussion, not just passive listening. Pupils reflect on the presentation and the information they have learned. Individual mastery is assessed.

Groups

For this study, groups were chosen at random to become the experiment and the control groups. In the experiment group, the science lesson was prepared according to the cooperative learning approach. In the control group, the science lesson was taught using a conventional training approach. The teacher of the experiment group was educated on how to treat science education with a cooperative learning approach before the lesson's application. Training was carried about by the same teacher in both groups. Since experimental research was treated, no population or sampling group process was stated.

For this research, participants were pupils of the seventh class studying at the Mimar Sinan Primary School in the city center of Eskişehir during the fall term of the 2005-2006 academic year. The research-developed achievement test and attitude scale were given to 7th grade pupils at Mimar Sinan Primary School. The pre-test and pre-attitude scores were evaluated according to the conclusions of a t-test. It was found that only two classes, 7-A and 7-C, were similar. The 7-A class was chosen as the experiment group and the 7-C class was chosen as the control group at random.

Data related to pre-test grades of the group. The results of the t-test determining the values of the pre-test grades, the standard deviation and whether there was a significant difference between the grades of the groups in which cooperative learning and conventional teaching were applied during the lesson on the structure of the atom in a primary school seventh class are given in Table 2. As seen in the table, the total mean value pre-test grades were 7,26 for the experiment group and 7,00 for the control group. To determine the significance of the difference between the mean values of the total of the seventh class pre-test grades for the experiment and control

groups, a t-test was applied using the SPSS statistical analysis program. The following data were provided at the end of the analysis. Related to the total mean values of the pre-test grades, the .555 t value was not found significant with a .658 p value, 52 degrees of freedom and a 0,05 significant level. Thus, the groups are equivalent to each other as to the mean values of the pre-test.

Table 2*Mean Values of Pre-Test Grades of the Groups*

Participants	n	\bar{X}	s.d.	t
Experiment Group	27	7,26	1,61	,555
Control Group	27	7,00	1,82	

*p<.05

The results of the t-test determining the values of pre-attitude grades, the standard deviation and whether there is a significant difference between the grades of the groups in which cooperative learning and conventional teaching were applied during the lesson on the structure of the atom in a primary school seventh class are given in Table 3.

Table 3*Mean Values of Pre-Attitude Grades toward the Science Lesson of the Groups*

Participants	n	\bar{X}	s.d.	t
Experiment Group	27	120,93	15,43	,580
Control Group	27	118,56	14,61	

*p<.05

As seen in Table 3, the total mean values of pre-attitude grades are 120,93 for the experiment group and 118,56 is the control group. To determine the significance of the difference between the mean values of the total pre-attitude grades for the experiment and control groups, a t-test was applied using the SPSS statistical analysis program. The following data were provided at the end of the analysis. As related to the total mean values of the pre-test grades, a .580 t value was not found significant with a .543 p value, 52 degrees of freedom and a 0,05 significant level. Thus, both groups are equivalent to each other as to the mean values of the pre-attitude grades.

Findings and Results

The Results Pertaining to the First Sub-problem and Discussion

The first sub-problem is, "Is there any significant difference between the mean of the achievement scores for a cooperative learning approach and the mean of achievement scores for a conventional training approach in the teaching of science?" To test the first sub-problem, the mean value of achievement (the difference between the pre-test and post-test) and the standard deviations of the experiment and control groups were calculated. The data may be seen in Table 4.

Table 4

Mean Values of Differences between Pre- and Post-Test Grades, Standard Deviations, and t Values of Experiment and Control Groups

Participants	number of questions	n	pre-test score \bar{X}	post-test score \bar{X}	difference between pre- and post-test \bar{X}	standard deviation s.d	t
Experimental Group	25	27	7,26	17,52	10,26	2,07	8,738***
Control Group	25	27	7,00	12,70	5,70	1,75	

***p<.001

As seen in Table 4, the total mean value difference between pre- and post-test grades is 10,26 for the experiment group and 5,70 is the control group. To determine the significance of the difference between the mean values of the difference between pre- and post-test grades for the experiment and control groups, a t test was applied using the SPSS statistical analysis program. The following data were provided at the end of the analysis. Related to the mean values of difference between pre- and post-test grades, an 8,738 t value was found significant with a .001 p value, 52 degrees of freedom and a 0,05 significant level. According to the data, there is a significant difference in the mean values of difference between the pre- and post-test grades of the experiment and control groups, and the cooperative learning approach is more efficient than the conventional approach.

The total mean values of post-test grades and standard deviations of both groups were also evaluated. To determine the significance of the difference between the mean values of the post-test grades for the experiment and control groups, a t test was applied using the SPSS statistical analysis program. The data is shown in Table 5.

Table 5

Mean Values of Post-test Grades, Standard Deviations and t Values of Experiment and Control Groups

Participants	n	post-test score \bar{X}	standard deviation s.d	t
Experiment Group	27	17,52	1,63	8,333 ***
Control Group	27	12,70	2,52	

***p<.001

As seen in Table 5, the total mean value of post-test grades are 17,52 for the experiment group and 12,70 is the control group. To determine the significance of the difference between the mean values of the post-test grades for the experiment and control groups, a t test was applied using the SPSS statistical analysis program. The following data were provided at the end of the analysis. Related to the mean values of the post-test grades, an 8,333 t value was found significant with a .001 p value, 52 degrees of freedom and a 0,05 significant level. According to the data, there is a significant difference in the mean values of difference between pre- and post-test grades in the experiment and control groups, and the cooperative learning approach is more efficient than the conventional approach.

The literature supports the findings of this research concerning this sub-problem. Arslan, Bora and Samancı (2006) researched the effects of cooperative learning techniques on learning about the nervous system with tenth class pupils. In this research, pupils participated actively, and it was concluded that there is a positive effect from the cooperative learning approach on pupil success. Doymuş and Şimşek (2007) researched the effects of the Jigsaw technique and pupils' opinions of this technique while learning chemical bonds. At the end of the research, the academic success of the class that used the Jigsaw technique improved significantly. Kınal, Ergül and Timur (2007) researched the effects of cooperative learning on pupils' success when learning science. They found a significant difference in favor of the cooperative learning approach. Şenol, Bal and Yıldırım (2007) researched the effects of cooperative learning on pupils' success while while studying sense organs in the sixth class. In this experiment, it was determined that the cooperative learning approach improved pupil success. In conclusion, science lessons taught using the cooperative learning approach increase pupil success.

Results Relating to the Second Sub-problem and Discussion

The second sub-problem is, "Is there any significant difference in the mean scores of the post-attitude grades about the science lesson between the cooperative learning approach group and the conventional learning approach group?" To test the second

sub-problem, the mean value of the post-attitude grades and standard deviations of the experiment and control groups were calculated. The data may be seen in Table 6.

Table 6

Mean Value of Post-Attitude Grades, Standard Deviation and t value of Experiment and Control Groups

Participants	n	post-attitude \bar{X}	standard deviation sd	t
Experiment Group	27	137,22	10,23	3,472 ***
Control Group	27	127,67	10,00	

***p<.001

As seen in Table 6, the total mean values of the post-attitude grades are 137,22 for the experiment group and 127,67 for the control group. To determine the significance of the difference between the mean values of the post-attitude grades for the experiment and control groups, a t test was applied using the SPSS statistical analysis program. The following data were provided at the end of the analysis. Related to the mean values of the post-attitude grades, a 3,472 t value was found significant with a .001 p value, 52 degrees of freedom and a 0,05 significant level. According to the data, there is a significant difference in the mean values of post-attitude grades for the experiment and control groups, and the cooperative learning approach is more efficient than the conventional approach.

Studies that support this finding exist in the literature. Altınok and Açıköz (2006) researched the effects of attitude on science lessons of cooperative and individual concept mapping. They concluded that pupils who used concept mapping in their science class developed positive attitudes about the science lesson. Şenol et al. (2007) researched the effects on pupil's attitude of a cooperative learning approach while teaching sense organs to a sixth class. They determined that the experiment group using the cooperative learning approach developed positive attitudes about the science lesson.

Conclusions and Recommendations

There is a significant difference in the mean value of the difference between pre- and post-test grades when comparing groups taught a science lesson using either the cooperative learning approach or a conventional learning approach. The cooperative learning approach is more efficient than the conventional approach. The pupils taught with a cooperative learning approach developed positive attitudes about the science lesson. As a result, the application of cooperative learning affected achievement and attitude. According to the results of this study, if the cooperative

learning approach is used in the teaching of science in seven grade classes, it can increase pupil success and improve attitude. The evidence of this study is consistent with the results of these researchers: Arslan et al., (2006), Altınok and Açıkgöz (2006), Demir (2008), Doymuş and Şimşek (2007), Doymuş, Şimşek and Karaçöp (2007), Doymuş et al., (2009), Kayıran and İflazoğlu (2007), Korkut-Owen, Owen and Ballesterio (2009), Kınal et al., (2007), and Şenol et al., (2007). On the other hand, the findings of this study partially conflict with Altınparmak and Nakipoğlu's (2005), Tanuşlı and Sağlam (2006), and Candaş Karababa's (2009) findings. Finally, the findings of the research and its results should be evaluated carefully. It should be kept in mind that this study was on only one subject of science; it did not include other subjects. Suggestions can be given based only on these results.

Sample training circumstances belonging to different lessons can be prepared which are appropriate for the cooperative learning approach. While evaluating the learning products, the process has to give special importance to the curricula to which the cooperative learning approach is administered. To do this, pupil participation in class activities and portfolios can be emphasized. Descriptive studies may be carried out to determine to what extent teachers have knowledge of cooperative learning and their willingness to use cooperative learning. Moreover, descriptive studies may also be conducted on the differences in the use of cooperative learning within specific countries and abroad.

References

- Altınok, H., & Açıkgöz, K. Ü. (2006). İşbirlikli ve bireysel kavram haritalamanın fen bilgisi dersine yönelik tutum üzerindeki etkileri [Effects of cooperative and individual concept mapping on attitudes toward science] *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi- Hacettepe University Journal of Education*, 30, 21-29.
- Altınparmak, M., & Nakipoğlu, M. (2005). Lise biyoloji laboratuvarlarında işbirlikli öğrenme yönteminin tutum ve başarıya etkisi [Cooperative learning approach in high school biology laboratory the influence on achievement and attitude]. *Gazi Üniversitesi Türk Eğitim Bilimleri Dergisi*, 1(3).
- Arslan, O., Bora, N. D., & Samancı, N. K. (2006). İşbirliğine dayalı öğrenme tekniklerinin 10. sınıf öğrencilerinin sinir sistemi konusunu öğrenmelerine etkisi [The effect of cooperative learning strategies on 10th grade pupils' achievement on nervous system]. *Eğitim Araştırmaları*, 6(23), 1-9.
- Candaş Karababa, Z.C. (2009). Effects of cooperative learning on prospective teachers' achievement and social interactions. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi- Hacettepe University Journal of Education*, 36, 32-40.
- Demir, K. (2008). Transformational leadership and collective efficacy: The moderating roles of collaborative culture and teachers' self-efficacy. *Eğitim Araştırmaları- Eurasian Journal of Educational Research*, 33, 93-112.
- Ding, M., Li, X., Piccolo, D. & Kulm, G. (2007). Teacher interventions in cooperative learning mathematics classes. *The Journal of Educational Research*, 100 (3), 162-176.

- Doymuş, K. & Şimşek, Ü. (2007). Kimyasal bağların öğretilmesinde jigsaw tekniğinin etkisi ve bu teknik hakkında öğrenci görüşleri [The effect of the teaching of chemical bonds jigsaw techniques and pupil opinions about this technique] *Milli Eğitim Üç Aylık Eğitim ve Sosyal Bilimler Dergisi*, 35 (173), 231-244.
- Doymuş, K., Şimşek, Ü., & Karaçöp, A. (2007). Genel kimya laboratuvarı dersinde öğrencilerin akademik başarısına, laboratuvar malzemelerini tanıma ve kullanımına işbirlikli ve geleneksel öğrenme yönteminin etkisi. [The effect of cooperative learning and traditional method on students' achievements, identifications and use of laboratory equipments in general chemistry laboratory course]. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 28, 31-43.
- Doymuş, K., Şimşek, Ü., & Karaçöp, A. (2009). The effects of computer animations and cooperative learning methods in micro, macro and symbolic level learning of states of matter. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 36, 109-128.
- Ghaith, G. (2002). The relationship between cooperative learning, perception of social support, and academic achievement. *American University of Beirut*, 30(3), 263-273.
- Gillies, R. M. (2004). The effects of communication training on teachers' and students' verbal behaviors during cooperative learning. *International Journal of Educational Research*, 41(3), 257-279.
- Jacobs, G. (2006). *Cooperative learning and second language teaching*, Cambridge University Press.
- Johnson, D., Johnson, R., & Smith, K. (1998). *Active learning: cooperation in the college classroom*. Retrieved February 8 2009 from <http://clte.asu.edu/active/baselemcooptms.pdf>.
- Jolliffe, W. (2007). *Cooperative Learning in the Classroom: Putting it into Practice*. Retrieved July 11, 2009 from <http://books.google.com.jm/books?id=YCHKT00-xa4C&pg=PA1&dq=cooperative+learning&lr=>
- Johnson, D.W. & Johnson, R. T. (1990). *Cooperative learning*, Edina, MN: Interaction Book Company.
- Johnson, D. W., Johnson R. T. & Holubee E. (1998). *Cooperation in the classroom*, Edina, MN: Interaction Book Company.
- Kayıran, B. K., İflazoğlu, A. (2007). Çoklu zeka kuramı destekli kubaşık öğrenme yönteminin türkçe dersine ilişkin tutuma ve okuduğunu anlama başarısına etkisi. [The effects of cooperative learning method supported by multiple intelligences theory on attitudes toward türkish language course and reading comprehension achievement]. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 29, 129-141.
- Kıncal, R. Y., Ergül, R. & Timur, S. (2007). Fen bilgisi öğretiminde işbirlikli öğrenme yönteminin öğrenci başarısına etkisi [Effect of cooperative learning method to

- pupil achievement in science teaching]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi-Hacettepe University Journal of Education*, 32, 156-163.
- Korkut-Owen, F., Owen, D., & Ballestero, V. (2009). Counselors and administrators: the collaborative alliance an three countries. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 36, 23-38.
- Lin, E. (2006). Cooperative learning in the science classroom: A new learning model for a new year. *The Science Teacher*, 73(5), 34-39.
- Mahrn, M. A. (2000). *The effect of using a suggested program on developing some of the composition writing skills of first year secondary school students through cooperative learning*. M.A Thesis. Assuit University.
- McLeish, K.(2009). *Attitude of Students Towards Cooperative Learning Methods at Knox Community College: A Descriptive Study*. Liberal Studies University, Jamaica.
- Millis, B.J. (1995). *Teaching improvement practices: Successful strategies for higher education*. Bolton: Anker.
- Stahl, R. (1994). *The Essential elements of cooperative learning in the classroom*. Retrieved July 11, 2009 from http://www.eric.ed.gov/ERICDocs/data_ericdocs2sql/content_storage_01/0000019b/80/15/a4/6f.pdf.
- Slavin, R.E. (1991). Synthesis of research on cooperative learning. *Educational Leadership*, 48(5), 71-82.
- Şenol, H., Bal, & Yıldırım, H.İ. (2007). İlköğretim 6. sınıf fen bilgisi dersinde duyu organları konusunun işlenmesinde işbirlikli öğrenme yönteminin öğrenci başarısı ve tutum üzerinde etkisi. [The effects of cooperative learning approach on pupils' success while teaching the unit called "sense organs" of sixth class of science lesson of primary training]. *Kastamonu Eğitim Dergisi*, 15(1).
- Tanışlı, D., Sağlam, M. (2006). Matematik öğretiminde işbirlikli öğrenmede bilgi değişme tekniğinin etkililiği. [Effectiveness of the exchange of knowledge method in cooperative learning in mathematics teaching]. *Eğitimde Kuram ve Uygulama*, 2(2), 47-67.
- Vygotsky, L.S. (1987). *Problems of general psychology including the volume thinking and speech*. New York: Plenum Press.
- Webb, J. (2002). *Benefits of cooperative learning in a multimedia environment*. M.A dissertation, B.S, Southern Illinois University.

Fen Bilgisi Öğretiminde İşbirlikçi Öğrenme Yaklaşımı

(Özet)

Problem Durumu : İşbirliğine dayalı öğrenme eğitim alanında son zamanlarda ortaya çıkmış yeni bir yaklaşım değildir. İşbirliğine dayalı öğrenmenin tarihi oldukça eskidir. Çok eskiden beri öğretmenler öğrencilerine grup projeleri, grup tartışmaları, grup çalışmaları yaptırmışlar ya da onları bu yönde desteklemişlerdir. Ancak yapılan her grup çalışmasını işbirliğine dayalı öğrenme olarak adlandırmak yanlış olur. İşbirlikçi öğrenme, öğrencilerin küçük gruplar oluşturarak bir problemi çözmek ya da bir görevi yerine getirmek üzere ortak bir amaç uğruna birlikte çalışma yoluyla bir konuyu öğrenme yaklaşımıdır. İşbirlikçi öğrenme gruplarında öğrenciler grup arkadaşlarının öğrenmesini en üst seviyeye çıkarmaya çalışırlar. Grup arkadaşları arasındaki iyi ilişkileri korurlar. Pozitif bağımlılık, grup ödülü, yüz-yüze etkileşim, bireysel değerlendirilebilirlik, sosyal beceriler, grup sürecinin değerlendirilmesi, eşit başarı fırsatı işbirlikçi öğrenme için gerekli olan koşullar ve özelliklerdir. İşbirlikçi öğrenmede her bir üyenin başarısı ya da başarısızlığı grupta yer alan bütün bireyler tarafından paylaşılır. İşbirliği yapan gruplarda üyeler arasında pozitif bir dayanışma bulunur. Grup üyelerinin üzerlerine aldıkları ve bireysel olarak yüklendikleri sorumlulukları bulunur. Öğrenciler grup olarak değerlendirilmelerinin yanında bireysel olarak da değerlendirilirler. Grup üyeleri yetenek, cinsiyet, başarı ve kişisel özellikleri bakımından heterojen olarak belirlenir. Gruplardaki bütün üyeler liderlik görevini paylaşırlar. Sosyal ve kişilerarası ilişkilerdeki beceriler ve beraber çalışmanın gerekliliği öğrencilere doğrudan öğretilir. Öğretmen gözlem yapar, grupta ortaya çıkan sorunları analiz eder ve gruplara görevlerini nasıl yapacaklarına dair rehberlik eder. Öğretmen gruplardaki işbirliği ve verimin artırılması için grupların oluşturulmasından grup üyelerinin değerlendirilmesine kadar ki bütün aşamaların planlanmasını yapar.

Genel olarak işbirlikçi öğrenme yaklaşımında öğretimin düzenlenmesi altı basamakta gerçekleştirilir. 1.Öğretmen tarafından belirlenen genel bir problem alanı içindeki konular öğrenciler tarafından seçilir, daha sonra gruplar oluşturulur. 2.Dersle ilgili alt konular, öğretim yöntemleri öğretmen ve öğrenciler tarafından işbirliğine dayalı olarak planlanır ve belirlenen problem konularına göre hedef davranışların düzeyleri belirlenir. 3.Öğrenciler, belirlenen hedef davranışlara göre kendi planlarını gerçekleştirirler. 4.Grup içindeki öğrenciler birlikte elde edilen bilgileri analiz ederler ve değerlendirirler. Daha sonra her bir grup üyesi kendine daha önceden verilen konuyu grup arkadaşlarına nasıl anlatması gerektiğini planlar. 5.Her bir gruptaki öğrenciler grup arkadaşlarının konuları anlamasını sağlar. Grup sunuları öğretmen tarafından düzenlenir. 6.Sınıftaki çalışmalar gerektiğinde öğretmen ve öğrenciler tarafından birlikte değerlendirilir ve bir karara bağlanır.

İşbirlikçi öğrenme yaklaşımı tek bir teknikten oluşmadığı için uygulanacak teknik dersin amacına, konuya, öğrencilerin öğrenme seviyelerine, yetenek ve eldeki olanaklara bakılarak önceden tespit edilir. Çalışma alanı ve sınıf öğrencilerin katılımını ve konsantrasyonlarını olumlu yönde etkileyecek biçimde düzenlenir. Gruplar öğrenciler belli bir seviyeye gelinceye kadar öğretmen tarafından belirlenir. Grup büyüklüğü var olan materyal, akademik işin türü, çalışmaya ayrılan zaman süreci, öğrencilerin sahip olduğu sosyal becerilere göre değişiklik gösterir. Çalışmanın akışını bozmadan çıkabilecek gürültüyü azaltmak için bazı önlemler alınır. Öğretmen uygulamayı başlatmadan önce akademik işi, görevi, hedefleri, kavram ve stratejileri, başarı için ölçütleri açıklayarak grup çalışmasını ve amacını öğrencilere açık bir şekilde ifade eder. Çalışmalar devam ederken öğretmen, fikir ve

düşüncelerin paylaşılmasında, gruptaki her öğrencinin katılımının sağlanmasında düşüncelerin birbirleriyle ilişkilendirilmesinde, grup raporunun veya projesinin hazırlanmasında destek olur. Bir grubun çalışmasını diğerlerinden önce bitirmesi durumunda grubun yaptıkları öğretmen tarafından kontrol edilir. Önce bitiren grubun öğrencileri diğer gruplara yardım için gönderilir. Gruplar yaptıkları çalışma ile ilgili bir sonuca ulaştığında çalışmalarını bitirirler. Değerlendirme işlemine öğrencilerde katılır, öğrenmenin nicelik ve niteliği belirlenir.

Öğrenme öğretme ortamında öğrencilerin derse aktif katılımlarını ve kalıcı öğrenmeyi sağlayacak öğrenme yaklaşımlarının kullanılması önemli bir boyuttur. Etkili öğrenme yaklaşımlarının kullanılması öğretimin kalitesini arttıracaktır. Çocuklar en iyi yaparak ve yaşayarak öğrenirler. Dersin daha etkili ve verimli olması için Fen Bilgisi öğretim programları öğrenci merkezli, araştırmaya, incelemeye ve deneye dayalı etkinliklerden oluşmalıdır. Başarı düzeylerini arttırmada geleneksel öğrenme yöntemlerine göre fark bulunan ve öğrenme öğretme süreçlerine karşı olumlu tutumlar geliştirmeyi sağlayan işbirliğine dayalı öğrenme yaklaşımının Fen Bilgisi dersine etkilerinin araştırılması faydalı olacaktır.

Araştırmanın Amacı: Bu araştırmanın amacı, Fen bilgisi dersinin öğretiminde işbirlikçi öğrenme yaklaşımının uygulandığı grubun başarı ve tutum puanlarının ortalamaları ile geleneksel öğretimin uygulandığı grubun başarı ve tutum puanlarının ortalamaları arasında anlamlı bir farkın olup olmadığını ortaya koymaktır.

Araştırmanın Yöntemi: Araştırmada deneysel yöntem kullanılmıştır. Bu yöntemin kontrol gruplu ön test, son test deseni uygulanmıştır. Araştırmada 2005-2006 öğretim yılı güz döneminde Eskişehir ili merkezinde bulunan Mimar Sinan İlköğretim okulu 7. sınıfı devam eden öğrencilerden yararlanılmıştır. Deney ve kontrol gruplarının ön test puanları ve ön tutum puanlarının ortalamalarına bakılarak birbirine benzeyen iki sınıf seçilmiştir. Bunlardan 7-A sınıfı deney ve 7-C sınıfı kontrol grubu olarak belirlenmiştir.

Araştırmanın Bulguları: Fen bilgisi dersinde işbirlikçi öğrenme yaklaşımının uygulandığı deney grubu ile geleneksel öğretimin uygulandığı kontrol grubunun başarı ve tutum puanlarının ortalamaları arasında anlamlı bir fark bulunmuştur. Bu fark, deney grubunun lehinedir. İşbirlikçi öğrenme yaklaşımının uygulamaları ile öğretim daha etkili olmuştur. Fen bilgisi dersinde işbirlikçi öğrenme yaklaşımının uygulandığı sınıftaki öğrenciler, fen bilgisi dersine yönelik olumlu tutum geliştirmişlerdir. Sonuç olarak, işbirlikçi öğrenme yaklaşımının uygulamaları, başarı ve tutum üzerinde etkili olmuştur.

Araştırmanın Önerileri: İşbirlikli öğrenme teknikleri fen bilgisi dersinin farklı konularında deneyerek araştırma sonuçları değerlendirilebilir. İşbirlikçi öğrenme tekniklerinin sınıf ortamında kullanımına ve süreç değerlendirmeye ağırlık verilebilir. Öğretmenlerin işbirlikçi öğrenme yaklaşımı hakkında ne kadar bilgi sahibi oldukları ve bunu uygulamak için ne kadar istekli oldukları üzerine betimsel araştırma yapılabilir. İşbirlikçi öğrenmenin yurtiçi ve yurtdışı uygulamaları arasındaki farklılıkları konu edinen betimsel çalışmalar yapılabilir.

Anahtar sözcükler: İşbirlikçi öğrenme yaklaşımı, fen bilgisi, başarı, geleneksel öğrenme yaklaşım, tutum.

Development and Application of Three-Tier Heat and Temperature Test: Sample of Bachelor and Graduate Students

Ali Eryılmaz*

Suggested Citation:

Eryılmaz, A. (2010). Development and application of three-tier heat and temperature test: Sample of bachelor and graduate students. *Eğitim Araştırmaları - Eurasian Journal of Educational Research*, 40, 53-76.

Abstract

Background of Study: Two-tier items compensate the limitation of simple multiple choice items that cannot measure the reason of selection of the alternatives. Three-tier items compensate the limitation of two-tier items that cannot measure the certainty of the responses. It is not common to use three-tier tests in the literature and there are validity and reliability problems with three-tier tests.

Purpose of Study: The aim of this study is twofold: i) to develop and validate a three-tier test to assess the misconceptions about heat and temperature, and ii) to compare Turkish bachelor, master, and PhD students' /graduates' misconceptions by using different methods including a new uncommon coding style.

Methods: "Heat and temperature are the same" misconception was selected from the literature because percentages of the students who had these misconceptions were high among Turkish and international students. A three-tier test was developed by student interviews, then open-ended questions and categorizing students' responses to these questions. The students were asked to compare heat and temperature of different size desks made of the same materials, and the same size desks made of different materials. Furthermore, the relationship between heat and temperature concepts in these contexts was asked. A total of 1,619 bachelor, 219 master, and 60 PhD students or graduates took the online test asynchronously. The participants' correct and misconception scores were calculated by using the first-tiers, first two-tiers and all three-tiers.

* Assist. Prof. Dr., Middle East Technical University, Faculty of Education, eryilmaz@metu.edu.tr

Findings and Results: Results showed that the percentage of the participants' misconceptions varied between 5% and 48% for bachelor students, between 4% and 49% for master students, and between 7% and 40% for PhD students depending on whether participants manifested the misconception in every situation or at least in one situation. The results were supported by reliability and factor analyses.

Conclusions and Recommendations: Simple multiple choice questions have limitations to assess students' achievement and misconceptions compared to three-tier multiple choice questions. False positives, false negatives, and lack of knowledge should be taken into account to assess students' achievement as well as misconceptions. A score for each misconception must be calculated instead of using a total test score. This must be supported by suitable reliability and factor analyses. The test given in the appendix could be used by teachers and/or researchers to better evaluate students' ideas about heat and temperature.

Keywords: Physics education, three-tier test, misconception test, heat, temperature

Acknowledgement: This article is based on a paper presented to the European Conference on Educational Research, September 2009, Vienna, Austria.

For three decades students' misconceptions in physics have been widely studied in the literature. It is claimed that students' misconceptions are one of the most important factors that affect students' achievement in physics (Halloun & Hestenes, 1985). Therefore, more recent research studies have been concentrating on how to assess students' misconceptions with a more reliable, valid, and feasible method (Caleon & Subramaniam, 2009; Chen, Lin, & Lin, 2002; Hasan, Bagayoko, & Kelley, 1999; Kılıç & Sağlam, 2009; Rebello & Zollman, 2004; Tan, Goh, Chia, & Treagust, 2002). Although most teachers and researchers are familiar with the development of achievement tests, they may not be familiar with that of misconception tests. Table 1 summarizes the differences between these two tests that could be found in related textbooks and literature. These comparisons are very important to consider when developing a new misconception test.

Table 1*Differences between Achievement and Misconception Tests*

Criteria	Achievement Tests	Misconception Tests
Purpose	Usually summative, and rarely formative and placement	Mostly Diagnostic
Type of Items	Conceptual and quantitative	Mostly conceptual
Scoring Given	To the correct answer	To the correct answer or alternative that reveals students' misconceptions
Table of Specification	Objectives categorized by a taxonomy	Misconception list
Evaluation	Mostly norm-referenced	Mostly criterion-referenced
Reliability	Usually higher than misconception tests	Usually less than achievement tests

The most popular misconception test in physics education has been the Force Concept Inventory (FCI) (Hestenes, Wells, & Swackhamer, 1992). Since it was first published as Mechanics Diagnostic Test (Halloun & Hestenes, 1985), it has been the most accepted and widely used test in the literature (Hake, 1998). However, it was also criticized in the literature in terms of what it really measures (Heller & Huffman, 1995; Huffman & Heller, 1995), percentages of false positives (correct answer with wrong reason) and false negatives (wrong answer with correct reason), and inadequacy of alternatives of the items (Rebello & Zollman, 2004). Critiques in the first part had been focused on the mismatch between the proposed dimensions and the factor analysis results of the FCI.

Although first Hestenes and Halloun (1995) used the concepts of false positives and false negatives to show content validity of the FCI results in response to Huffman and Heller (1995), critiques in the second part could be more objective if the FCI had a mechanism of calculating percentages of false negatives and false positives. For this purpose, two-tier test items have been used in the literature (Chen et al., 2002; Rollnick & Mahooana, 1999; Tan et al., 2002). In the first-tier, concept or the problem is asked. In this sense, the FCI items could be first-tiers of a two-tier test. In the second-tier, the reason or justification of the answer given to the first-tier is asked. In order to determine lack of knowledge, some researchers have used the third-tier (Caleon, & Subramaniam, 2009; Hasan et al., 1999). In the third-tier, students are asked to what extent they are sure about their responses. Table 2 helps us score a three-tier question. If you want to use a three-tier test to assess students' conceptual understanding as an achievement test, you should give a score to the

students who choose the correct alternative in the first-tier and the corresponding correct reason in the second-tier and they must be sure about these answers. Likewise, if you want to use a three-tier test to assess students' misconceptions as a diagnostic test, you should give a score to the students who choose the misconception alternative in the first-tier and the corresponding reason in the second-tier and they must be also sure about these answers. From Table 2, the other combinations are false negatives, false positives and lack of knowledge. When participants are not sure about their responses, all combinations of the correct and wrong responses of the first and second-tiers should be categorized as lack of knowledge even if both responses are correct. Thus, using simple multiple choice items, both correct and misconception scores are overemphasized.

Table 2

Decision Table Based on Combinations of the Results of the First and Second-tiers When Participant is Sure about her/his Responses.

		Results of the Second - Tier	
		Correct	Wrong
Results of the First - Tier	Correct	Knowledge of correct concept	Mostly false positive Rarely misconception
	Wrong	False negative	Mostly misconception about concept Rarely mistake

For the last critique, after administering four questions of the FCI and equivalent essay type form of the same questions to the bachelor students, Rebello and Zollman (2004) concluded that although the percentages of the correct responses were not changed significantly in each of two formats, the items in the FCI do not have the distracters that students responded to in the essay type formats. Thus distracters of the items in the FCI did not suffice for all students. In this regard, last criticism could be a problem for all selection type (true-false or alternative response, matching, and multiple-choice) items that require the students to select the answer. By this way, students' responses are restricted. Hence alternatives must be changed according to the physics backgrounds of the students. In order to solve this problem, a blank alternative could be included with the multiple-choice items in case students want to write an answer that is not provided.

Like the FCI, the thermal concept evaluation test was developed to assess students' understanding of thermodynamics concepts by Yeo and Zadnik (2001). It is thus most probable that it has the same limitations. Another limitation for both tests is that one cannot calculate a score for a specific misconception because of the number of the questions assessing each misconception. Although both tests have six items for some misconceptions, they have only one or two items for most

misconceptions (23 out of 35 for the thermal concept evaluation test and 16 out of 30 for the FCI). These items for each misconception are used while calculating percentages of students who hold that misconception. If each misconception were assessed by more than two items, then test length could be a problem. For this reason, one of the most common misconceptions about thermodynamic concepts was chosen for this study: "Heat and temperature are the same" or "temperature is a measure of heat" or "heat and the temperature are not the same but they can be used interchangeably in the context".

In light of the debated issues, the purposes of this study are i) to develop and validate a three-tier test to assess the misconceptions about heat and temperature, and ii) to compare bachelor, master, and PhD students' misconceptions by using a new uncommon coding style and old styles.

Method

Development and Administration of the Instrument

First, the related literature was examined for misconceptions about heat and temperature and for popularity of these misconceptions among students. The misconceptions in Appendix A were selected to be studied because percentages of the students who had these misconceptions were high among Turkish (Başer, 1996) and international (Yeo & Zadnik, 2001) students. Then, interview questions were constructed. Two high, two medium, and two low ability students (as judged by the physics teacher of the classroom) were interviewed with these questions. The purpose of the interviews was to make sure that the questions are clear, understandable, and help elicit students' misconceptions under investigation. Next, revised open-ended questions were prepared and administered to 36 high school students. Furthermore, students' responses were categorized, and three-tier questions were prepared. This test was checked by several physics education PhD students and two university physics instructors to ensure that the items were properly constructed and relevant to the misconceptions, and correct and misconception codings given in Appendix A were correct. The English version of final form of the three-tier heat and temperature test is given in Appendix B. Next, the final form of the test was published on a website. The possible visitors (members of the website and related discussion forms) were informed by e-mail. The visitors completed the test at their convenience. Therefore, there is no time limitation to finish the test.

The website is a physics and physics education portal that has been on the Internet since 2003, and accessed daily by approximately 4,000 visitors. There are biographies of the scientists, simulations, tutorials, answers to the physics questions of the university entrance examinations, and forums about different topics related to physics. There are approximately 22,000 registered members of the website.

Subjects

The participants of the study were the individuals who retrieved the website and asynchronously took the test. The test was completed by 3,879 times in two and a half months. Four hundred seventy-four multiple entries were found by controlling

the responses that they selected or wrote. Thus, 3,405 entries were retained. There were four demographic questions in the beginning of the test: "What is your gender?", "What is your educational level?", "What is your occupation?", and "When were you born?" One thousand five hundred and seven of the participants chose that they are elementary or high school students/graduates. These students' responses were not discussed here. The remaining participants' demographics are given in Table 3. When the participants chose bachelor's degree as the educational level, bachelor could mean that they are bachelor students or they hold a bachelor's degree. The same could be valid for other educational levels. Unlike birth year and occupation, there were no missing data for education level and gender.

Table 3

Participants' Educational Level, Gender, Birth Date and Occupations

Education Level	Gender	Birth Year	Occupation
1,619 Bachelor	1,130 Males	Max = 1988	866 Students
	489 Females	Min = 1940 Median = 1982	391 Teachers 124 Engineers 24 Officers 20 Self-employed 14 Unemployed 160 Others
		N = 1,431	N = 1,599
219 Master	134 Males	Max = 1995	96 Teachers
	85 Females	Min = 1950 Median = 1980	51 Students 17 Research Assistants 8 Engineers 36 Others
		N = 193	N = 208
60 PhD	40 Males	Max = 1981	21 Academicians
	20 Females	Min = 1943 Median = 1974	12 Students 9 Physicians 6 Teachers 9 Others
		N = 50	N = 57

If possible, reformat the second row so that N=193 is on the same level as N=208.

Also, can the spacing be reduced between 17 and Research Assistants?

Findings and Results

The data were taken from the website as an electronic document. Correct and misconception scores were calculated by MS-Excel according to the coding tables given in Appendix A. Reliability and factor analyses of the data were done by Statistical Packages for the Social Sciences (SPSS).

Results of the Blank Alternatives

From Table 4, only 3.3% of the participants filled the blank alternative for each question. As the education level increases, the number of the participants who filled the blank alternatives slightly increases. Compared to the other questions, the fewer participants filled the blank alternatives of the first tiers of the questions asking temperature and heat of the same material desks with different sizes (Temp1.1 and Heat1.1). For these questions, participants either wrote the correct answers in their words or asked for extra information (temperature of the room, initial temperature of the desks, and mass and surface areas of the desks). For the first tiers of the questions asking temperature and heat of the same size desks with different materials (Temp2.1 and Heat2.1), most of the participants who filled the blank alternatives asked for extra information about temperature of the room, initial temperature of the desks, and mass and specific heat capacities of the iron and wood desks.

Table 4

Number of the Participants who Filled the Blank Alternative for Each Question

Question	Bachelor	Master	PhD	Total
Temp1.1	8	1	0	9
Temp1.2	85	17	3	105
Temp2.1	19	2	0	21
Temp2.2	74	12	4	90
Heat1.1	9	3	3	15
Heat1.2	68	5	1	74
Heat2.1	62	13	3	78
Heat2.2	61	10	5	76
HeatTemp	70	16	5	91
Average %	3.1	4.0	4.4	3.3

For Temp1.2 and Temp2.2 questions, approximately 75% of the participants who filled the blank alternatives wrote either because the desks reached the thermal equilibrium or because the desks have been resting in the room for a long time. Ten participants for Temp1.2 and 13 participants for Temp 2.2 questions asked for extra

information (temperature of the room, mass and surface areas of the desks for Temp1.2 question, and temperature of the room, heat and specific heat capacity of the desks, and thermal conductivity of the desks for Temp2.2 question). The others wrote sentences that have the same meaning with Alternative "A" of the questions.

Most of the participants who filled the blank alternatives for Heat1.2 and Heat2.2 questions wrote sentences that have the same meaning with Alternative "a" of the questions. The others asked for some extra information related to the temperature of the room and specific heat capacities of the desks. For the last question, some participants who filled the blank alternative wrote sentences that have the same meaning with the alternatives (14 for the correct and 10 for "B" or "C" alternatives). Twenty one participants wrote sentences that include wrong definitions of both heat and temperature (For example, "heat is the total energy of a matter and temperature is an average energy of a matter", "heat is the kinetic energy of a matter and temperature is the average kinetic energy of a matter"). Seventeen participants wrote that temperature and heat are different without writing definitions. Correct definition of temperature but wrong definition of heat was written by 18 participants (For example, "temperature is a measure of the average kinetic energy of the individual atoms or molecules composing a substance", "heat is the total temperature of the matter", "heat of a matter depends on mass, temperature, and type of the matter", "heat is a function of temperature"). The others wrote something irrelevant in the blank alternative. In this way, almost all the sentences written in the blank alternatives were coded into the existing alternatives of the related questions.

Correct Scores

Table 5 shows the percentages of the participants who correctly answered the questions as simple multiple choice test, two-tier test, and three-tier test. The percentages of the participants who answered the first and second-tiers correctly are approximately the same. Moreover, the percentage of the participants who correctly compared the temperature of the desks is higher than that of those who correctly interpreted heat of the desks. Although the percentage of the participants who correctly interpreted the heat of the desks is low, the percentage of the participants who correctly interpreted the relationship between heat and temperature in the context is high.

When the questions are analyzed as two-tier test, 69% of the bachelor students/graduates gave the correct responses to both the first and second-tiers of the first question about temperature. Although 69% of the bachelor students/graduates correctly compared the temperature of the same material with different size desks, only 40% of the bachelor students/graduates correctly compared the temperature of the same size with different material desks. Furthermore, the percentages are decreased approximately 8 points from the results of the first-tiers only to the first two-tiers. In other words, approximately 8% of the participants did not correctly explain the reasons for their answers to the first-tiers.

When the correct scores were calculated by using all three-tiers, the percentages were decreased approximately 4% compared to the percentages due to two-tier test. In other words, 4% of the participants, who answered the first two-tiers correctly, were not sure about their answers.

Table 5
Correct Score Percentages of the Questions of the Participants

	Bachelor %	Master %	PhD %
Simple multiple-choice			
Temp1.1	77	78	73
Temp1.2	72	71	70
Temp2.1	47	53	62
Temp2.2	44	49	57
Heat1.1	6	9	27
Heat1.2	6	10	30
Heat2.1	15	16	25
Heat2.2	7	10	30
HeatTemp	59	58	58
First two-tiers			
Temp1	69	68	65
Temp2	40	45	52
Heat1	4	7	23
Heat2	6	6	20
All three-tiers			
Temp1	65	63	63
Temp2	37	42	50
Heat1	4	6	23
Heat2	5	5	20
HeatTemp	54	56	55

As aforementioned, three-tier questions enable us to calculate the percentages of false positives, false negatives, and lack of knowledge of the participants that are given in Table 6. All the values in general are low. These low values are indicators for high content validity of the test results (Hestenes & Halloun, 1995). The average percentages of the false positives are approximately the same for the bachelor, master, and PhD students/graduates. The average percentage of the PhD students/graduates who gave the wrong answer with correct reason is two to three

percent greater than that of the bachelor and master students/graduates. Furthermore, no differences in the average percentages of the lack of knowledge of the participants were observed.

Table 6

Percentages of the False Positives, False Negatives and Lack of Knowledge of the Participants

	Temp1	Temp2	Heat1	Heat2	Average
	%	%	%	%	%
Bachelor					
False Positive	7	6	2	8	5.8
False Negative	3	3	1	1	2.0
Lack of Knowledge	9	11	9	14	10.8
Master					
False Positive	10	7	2	8	6.8
False Negative	3	4	2	3	3.0
Lack of Knowledge	8	9	10	12	9.8
PhD					
False Positive	8	10	3	5	6.5
False Negative	5	5	7	10	6.8
Lack of Knowledge	7	12	13	13	11.3

Misconception Scores

From Table 7, the participants have fewer misconceptions about temperature of the desks rather than heat of the desks. In other words, they have more difficulty in understanding heat than temperature. Although the participants thought that the temperature of the desks depends mostly on the material the desks are made of, they thought that the heat of the desks depends mostly on sizes of the desks. By integrating the results presented in Table 5 and Table 7, one cannot say that if the participants do not know the answer to the question, they have a misconception. For example, although 65% of the bachelor students/graduates given in Table 5 correctly compared the temperature of the same material but small and large desks, only 13% of those given in Table 7 hold the misconception that the temperature of objects depends on their sizes even though they rest in the same room for a long time. The remaining 22% of those gave either the correct answer with wrong reason or wrong answer with correct reason or correct answer with correct reason but unsure about their answers. This percentage also includes wrong answer with wrong reason but the answer and corresponding reason are not related to this misconception. Furthermore, in general, the percentage of misconceptions is reduced more than 10% when the misconceptions are calculated by using all three-tiers compared to

calculations by using only first-tiers. These participants either made a mistake or had lack of knowledge. It is also noteworthy from Table 7 that from 47% to 80% of the participants have misconceptions about heat even with the three-tier test.

Table 7

Percentages of the Misconceptions According to the First, First Two and All Three-tiers

	Bachelor %	Master %	PhD %
First Tiers			
MisTemp1*	17	18	17
MisTemp2*	46	42	30
MisHeat1*	93	89	63
MisHeat2*	80	78	67
First two-tiers			
MisTemp1**	16	15	10
MisTemp2**	42	37	25
MisHeat1**	87	85	62
MisHeat2**	77	72	57
All three-tiers			
MisTemp1***	13	14	8
MisTemp2***	37	34	20
MisHeat1***	80	79	50
MisHeat2***	67	64	47

* First tiers were coded according to the coding table given in Appendix A.

** First two-tiers were coded according to the coding table given in Appendix A.

*** All three-tiers were coded according to the coding table given in Appendix A.

Table 8 shows participants' misconceptions according to the three-tier questions with different coding. This coding, given in Appendix A under the title of "Heat and temperature are the same", integrates the participants' responses about heat and temperature of the objects. The first-tiers of the temperature and heat questions for each context were coded as a question and the second-tiers of the temperature and heat questions were coded as another question, and the third-tiers of the temperature and heat questions were coded as another question. The coding checks whether the participants chose the same alternative for heat and temperature of the desks in the first, second, and third tiers of the questions. For example, 16% of the bachelor students/graduates chose the same alternative for the heat and temperature of the same material but small and large desks in the first tiers of the first questions. Then, 20% of the bachelor students/graduates chose the same alternatives for the heat and

temperature of the same material but small and large desks in the second tiers of the first questions. Furthermore, 13% of the bachelor students/graduates chose the same alternatives for the heat and temperature of the same material but small and large desks in the first and second tiers of the first questions. This value was reduced to 10% when all three-tiers were taken into account. The participants holding different degrees used the same sentence for heat and temperature of the same size desks more than the same material desks. The participants' misconceptions change from 5% to 48% for the bachelor, 4% to 49% for the master and 7% to 40% for the PhD students/graduates depending on whether participants manifested the misconception in every situation (AllAND) or at least in one situation (AllOR).

Table 8

Percentages of the Misconceptions According to the New Coding

	Bachelor %	Master %	PhD %
Question 1: The same material desks with different sizes			
First-tiers	16	16	18
Second-tiers	20	26	23
First two-tiers	13	15	15
All three-tiers	10	12	10
Question 2: The same size desks with different materials			
First-tiers	34	34	27
Second-tiers	41	41	30
First two-tiers	29	30	23
All three-tiers	24	25	20
All Questions			
AllOR*	48	49	40
AllAND**	5	4	7

*AllOR means all three-tiers of the 1st question OR all three-tiers of the 2nd question OR all tiers of the 3rd question (coding given in Appendix A).

**AllAND means all tiers of the 1st question AND all tiers of the 2nd question AND all tiers of the 3rd question (coding given in Appendix A).

Factor and Reliability Analyses of the Correct and Misconception Scores

Table 9 presents the reliability and factor analyses of the correct scores. Cronbach alpha reliability that assesses internal consistency of the test results and a principle component factor analysis with varimax factor rotation were performed by using SPSS. Although percentages of the correct scores were low especially for the heat questions, factor analysis results show that the participants' responses discriminate

heat and temperature concepts. The participants sometimes responded differently for the temperature of the same size but different material desks, and the same material

Table 9

Results of the Reliability and Factor Analyses of the Correct Scores

	Bachelor	Master	PhD
9 Questions*	3Factors - 67% - .63 1. H (.75) 2. T2 (.89) 3. T1+3Q (.60)	3Factors - 68% - .68 1. H (.79) 2. T2 (.88) 3. T1+3Q (.56)	2Factors - 68% - .80 1. H (.91) 2. T+3Q (.80)
8 Questions*	3Factors - 75% - .65 1. H (.75) 2. T2 (.89) 3. T1 (.84)	3Factors - 75% - .69 1. H (.79) 2. T2 (.88) 3. T1 (.80)	2Factors - 73% - .79 1. H (.91) 2. T (.83)
5 Questions** (First-tiers)	2Factors - 57% - .37 1. T+3Q (.47) 2. H (.50)	2Factors - 59% - .47 1. T (.59) 2. H+3Q (.41)	2Factors - 72% - .67 1. H (.88) 2. T+3Q (.63)
4 Questions** (First-tiers)	2Factors - 70% - .35 1. T (.57) 2. H (.50)	2Factors - 71% - .43 1. T (.59) 2. H (.54)	2Factors - 83% - .63 1. H (.88) 2. T (.69)
5 Questions*** (First Two-tiers)	2Factors - 63% - .39 1. H (.70) 2. T+3Q (.49)	2Factors - 65% - .48 1. H (.78) 2. T+3Q (.48)	2Factors - 77% - .69 1. H (.95) 2. T+3Q (.71)
4 Questions*** (First Two-tiers)	2Factors - 76% - .38 1. H (.70) 2. T (.60)	2Factors - 80% - .48 1. H (.78) 2. T (.62)	2Factors - 87% - .61 1. H (.95) 2. T (.71)
4 Questions**** (All Three-tiers)	2Factors - 75% - .43 1. H (.74) 2. T (.64)	2Factors - 77% - .53 1. H (.83) 2. T (.64)	2Factors - 87% - .63 1. H (.95) 2. T (.71)

* T1 covers temperature questions of Questions 1.1 and 1.2, T2 covers temperature questions of Questions 2.1 and 2.2, T covers all temperature questions, H covers all heat questions, and 3Q means Question 3.1.

** T covers temperature questions of Questions 1.1 and 2.1, H covers heat questions of Questions 1.1 and 2.1, and 3Q means Question 3.1.

*** T covers two temperature questions but Questions 1.1 and 1.2 coded as one question and Questions 2.1 and 2.2 coded as one question, H covers two heat questions but Questions 1.1 and 1.2 coded as one question and Questions 2.1 and 2.2 coded as one question, and 3Q means Question 3.1.

**** T covers two temperature questions but Questions 1.1, 1.2, and 1.3 coded as one question and Questions 2.1, 2.2, and 2.3 coded as one question, H covers two heat questions but Questions 1.1, 1.2, and 1.3 coded as one question and Questions 2.1, 2.2, and 2.3 coded as one question.

but different size desks. In other words, the participants think that the desks have the same temperature when they are made of the same material but their sizes are different. However they did not respond the same way when their sizes are equal but made of different materials. For example, reliability and factor analyses with the nine questions on the bachelor students/graduates yield three factors, and reliability of .63. Three factors explained 67% of the whole variance. Four heat questions (H) loaded in the first factor, two temperature questions about the same size but different material (T2) loaded on the second factor, remaining two temperature questions (T1) and the last question (3Q) loaded on the third factor. The Cronbach alpha reliabilities of the factors are .75, .89, and .60, respectively.

Table 10 shows the reliability and factor analyses of the misconceptions scores. Acceptable factors and high reliability coefficients were expected because participants' misconceptions were so high (Hestenes & Halloun, 1995). In other words, the participants could be taken as novices rather than an expert group about the concepts of heat and temperature. Reliability coefficients for the test results and for the factors varied between .46 and .87, and between .41 and .94, respectively. The misconceptions about heat and temperature loaded on the different factors except once for the PhD students/graduates when analyzing four misconceptions (first tiers). When misconceptions about the relationship between heat and temperature were included in the factor analyses, most of the time (seven out of nine) this misconception loaded on the same factor with the misconceptions about temperature. In other words, participants who thought that the temperature of the desks resting in a room for a long time depends on their sizes and/or materials, also thought that "heat and temperature are the same" or "temperature is a measure of heat" or "heat and the temperature are not the same but they can be used interchangeably in the context". The factors explained between 57% and 79% of the total variances and these values could be acceptable.

The first-tiers of the temperature and heat questions for each context were coded as a question and the second-tiers of the temperature and heat questions were coded as another question. Thus new coding formed four questions. By using this new coding, item analysis was performed. Results of the PhD students/graduates only were discussed here as an example. Difficulty indices for the PhD students/graduates ranged from .13 to .23 with a mean of .19. This means that the items were difficult for even the PhD students/graduates. Discrimination indices ranged from .60 to .79 with a mean of .70. These values could be accepted as quite high for a misconception test. Table 11 shows the reliability and factor analyses of the misconceptions scores with this new coding and the last question. The reliability coefficients varied between .69 and .86. First and second-tiers of a question, and first and third questions loaded on the same factor in all factor analyses. A unique factor was formed for the PhD students/graduates. This means that these participants'

responses about the relationship between heat and temperature did not change according to the context.

Table 10

Results of the Reliability and Factor Analyses of the Misconception Scores

	Bachelor	Master	PhD
9 Questions*	3Factors - 68% - .70	3Factors - 69% - .75	2Factors - 71% - .87
	1. H (.75)	1. H+3Q (.69)	1. H (.94)
	2. T1+3Q (.63)	2. T2 (.89)	2. T+3Q (.81)
	2. T2 (.78)	3. T1 (.90)	
8 Questions*	3Factors - 76% - .71	3Factors - 77% - .76	2Factors - 75% - .87
	1. H (.75)	1. H (.78)	1. H (.94)
	2. T1 (.93)	2. T2 (.89)	2. T (.80)
	3. T2 (.78)	3. T1 (.90)	
5 Questions** (First-tiers)	2Factors - 57% - .46	2Factors - 60% - .55	2Factors - 71% - .76
	1. T+3Q (.45)	1. H+3Q (.47)	1. H (.88)
	2. H (.52)	2. T (.60)	2. T+3Q (.64)
4 Questions** (First-tiers)	2Factors - 69% - .46	2Factors - 71% - .54	1Factor - 57% - .74
	1. H (.52)	1. T (.60)	1. H+T (.74)
	2. T (.52)	2. H (.54)	
5 Questions**** (All Three-tiers)	2Factors - 57% - .47	2Factors - 58% - .54	2Factors - 66% - .59
	1. T+3Q (.43)	1. T+3Q (.41)	1. H (.80)
	2. H (.58)	2. H (.59)	2. T+3Q (.51)
4 Questions**** (All Three-tiers)	2Factors - 69% - .48	2Factors - 72% - .59	2Factors - 79% - .57
	1. H (.58)	1. T (.58)	1. H (.80)
	2. T (.50)	2. H (.59)	2. T (.60)

*, **, **** Notations used in this table are the same as the corresponding notations under Table 9.

Table 11

Results of the Factor and Reliability Analyses of the Misconception Scores According to the New Coding.

	Bachelor	Master	PhD
5 Questions*	2Factors - 70% - .69	2Factors - 70% - .69	1Factor - 59% - .80
	1. 1QT+3QT (.54)	1. 2QT (.85)	1. 1QT+2QT+3QT (.80)
	2. 2QT (.84)	2. 1QT+3QT (.58)	
4 Questions*	2Factors - 86% - .77	2Factors - 85% - .76	1Factor - 70% - .86
	1. 2QT (.84)	1. 2QT (.85)	1. 1QT+2QT (.86)
	2. 1QT (.82)	2. 1QT (.79)	

* 1QT, 2QT, and 3QT were calculated for those students who are sure about their answers. 1QT stands for two questions: one for first-tiers of the first questions and one for second-tiers of the first questions, 2QT stands for two questions: one for first-tiers of the second questions and one for second-tiers of the second questions, and 3QT stands for the first-tiers of Question 3.

Implications and Suggestions

This study revealed the development and validation of the three-tier heat and temperature test. The study also revealed how to calculate correct and misconception scores from a three-tier test. The participants' correct and misconception scores were calculated by using the first-tiers, first two-tiers and all three-tiers. These helped us show the superiority of using three-tier tests to assess students' achievement and/or misconception. These scores were supported by the suitable reliability and factor analyses. This study suggests that;

1. Blank alternatives included with the multiple-choice items work fine with three-tier items. This could be taken as good practice with multiple choice items.
2. The test can be used to assess bachelor and graduate students' misconceptions about heat and temperature. By using three-tier tests, 13% of the bachelor students/graduates, 14% of the master students/graduates, and 8% of the PhD students/graduates certainly think that the temperature of the same

material desks but different sizes depends on the sizes of the desks even if the desks rest in a room for a long time. These values are 80%, 79%, and 50% for heat of the tables, respectively. In addition, 37% of the bachelor students/graduates, 34% of the master students/graduates, and 20% of the PhD students/graduates certainly think that the temperature of the same size desks but different materials depends on the material of the desks even if the desks rest in a room for a long time. These values are 67%, 64%, and 47% for heat of the tables, respectively. Furthermore, 32% of the bachelor students/graduates, 32% of the master students/graduates, and 30% of the PhD students/graduates certainly chose that "Heat and temperature are the same" or "Temperature is a measure of heat" or "Heat and temperature are not the same. However, they can be used interchangeably in the context".

3. Simple multiple choice questions have limitations to assess students' achievement and misconceptions compared to three-tier multiple choice questions. False positives, false negatives, and lack of knowledge should be taken into account to assess students' achievement as well as misconceptions. Rollnick and Mahooana (1999) effectively showed the speedy development of two-tier multiple choice diagnostic questions from simple multiple choice questions.
4. Misconception tests should not be scored based on the correct scores. Using correct scores on the misconception tests presumes that high correct scores on the test indicate few misconceptions, or low correct scores on the test indicate many misconceptions. In other words, if participants do not choose the correct alternative of a question about a concept, they have misconception about the concept. This study showed that this is not true even if the distracters are arranged in such a way that they reveal students' common sense.
5. A score for each misconception must be calculated instead of using total test score. This must be supported by suitable reliability and factor analyses. This may reduce the wrong use of total test score without verifying underlying unique dimension of the test. In this way, researchers can investigate the effectiveness of a treatment on a specific misconception. In this situation, number of questions to assess each misconception is an important factor to change the reliability of the test results.
6. Third-tiers of the test could be analyzed for a whole class to decide the modification of the instructional deliveries as used by Hasan et al. (1999). However, in this study, third-tiers were analyzed for each student to calculate the percentages of the lack of knowledge.
7. The test in Appendix B could be used by teachers and/or researchers to better evaluate participants' ideas about heat and temperature. The proposed format of the test and new coding could be used for other related but confusing concepts like position and velocity, velocity and acceleration, momentum and impulse, and so on.

References

- Başer, M. (1996). *Effect of conceptual change instruction on understanding of heat and temperature concepts and science*. Unpublished master thesis, Middle East Technical University, Ankara.
- Caleon, I., & Subramaniam, R. (2009). Development and Application of a Three-Tier Diagnostic Test to Assess Secondary Students' Understanding of Waves, *International Journal of Science Education*, Online first: 1-23.
- Chen, C., Lin, H., & Lin, M. (2002). Developing a two-tier diagnostic instrument to assess high school students' understanding - The formation of images by a plane mirror, *Proceedings of the National Science Council*, 12(3), 106-121.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *American Journal of Physics*, 66, 64-74.
- Halloun, I. A. & Hestenes, D. (1985). The initial knowledge state of college physics students, *American Journal of Physics*, 53(11), 1043-1055.
- Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the certainty of response index (CRI), *Physics Education*, 34(5), 294-299.
- Heller, P. & Huffman, D. (1995). Interpreting the force concept inventory: A reply to Hestenes and Halloun, *The Physics Teacher*, 33, 503-511.
- Hestenes, D. & Halloun, I. A. (1995). Interpreting the force concept inventory: A response to Huffman and Heller, *The Physics Teacher*, 33, 502-506.
- Hestenes, D. & Wells, M. & Swackhamer, G. (1992). Force concept inventory, *The Physics Teacher*, 30, 141-158.
- Huffman, D. & Heller, P. (1995). What does the force concept inventory actually measure?, *The Physics Teacher*, 33, 138- 143
- Kılıç, D. & Sağlam, N. (2009). Development of a two-tier diagnostic test to determine students' understanding of concepts in genetics, *Eurasian Journal of Educational Research*, 36, 227-244.
- Rebello, N. S. & Zollman, D. A., (2004). The effect of distracters on student performance on the force concept inventory, *American Journal of Physics*, 72(1), 116-125.
- Rollnick, M. & Mahooana, P. P. (1999). A quick and effective way of diagnosing student difficulties: Two-tier from simple multiple choice questions. *South African Journal of Chemistry*, 52(4), 161-164.

Tan, K. C. D., Goh, N. K., Chia, L. S., & Treagust, D. F. (2002). Development and application of a two-tier diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *Journal of Research in Science Teaching*, 39(4), 283-301.

Yeo, S. & Zadnik, M. (2001). Introductory thermal concept evaluation: assessing students' understanding, *The Physics Teacher*, 39, 496-504.

Appendix A: Coding Tables for the Correct and Misconception Scores.

Correct Scores:

Question No	Temperature	Heat
1.1.	C	d
1.2.	B	c
1.3.	A	a
2.1.	C	d
2.2.	B	c
2.3.	A	a
3.1.	D	
3.2.	A	

Misconception Scores:

MisHeatTemp: Heat and temperature are the same.

(3.1.A OR B OR C) AND (3.2.A)	OR	[(1.1.A AND 1.1.a) OR (1.1.B AND 1.1.b) OR (1.1.C AND 1.1.c) OR (1.1.D AND 1.1.d)]	OR	[(2.1.A AND 2.1.a) OR (2.1.B AND 2.1.b) OR (2.1.C AND 2.1.c) OR (2.1.D AND 2.1.d)]
	AND	AND [(1.2.A AND 1.2.a) OR (1.2.B AND 1.2.b) OR (1.2.C AND 1.2.c)]	AND	AND [(2.2.A AND 2.2.a) OR (2.2.B AND 2.2.b) OR (2.2.C AND 2.2.c)]
		AND (1.3. A AND 1.3. a)		AND (2.3. A AND 2.3. a)

MisTemp1: Temperature of an object depends on its size.

1.1.A OR B	AND	1.2.A	AND	1.3.A
------------	-----	-------	-----	-------

MisTemp2: Temperature of an object depends on the material made of. That is; iron is naturally cooler than wooden.

2.1.A OR B	AND	2.2.A	AND	2.3.A
------------	-----	-------	-----	-------

MisHeat1: "Heat of an object depends on its size" or "heat of two objects can be compared".

1.1.a OR b OR c	AND	1.2.a OR b	AND	1.3.a
-----------------	-----	------------	-----	-------

MisHeat2: "Heat of an object depends on the material made of" or "heat of two objects can be compared".

2.1.a OR b OR c	AND	2.2.a OR b	AND	2.3.a
-----------------	-----	------------	-----	-------

Appendix B: Three-tier Heat and Temperature Test.

1. Answer the following questions for two iron desk (one is small, one is large) resting in a room for a long time.	
1.1. Which one of the desk has higher temperature? A. Small desk has higher temperature. B. Large desk has higher temperature. C. Their temperature is the same. D. Their temperature cannot be compared. E. Other. Please write.....	1.1. Which one of the desk has higher heat? a. Small desk has higher heat. b. Large desk has higher heat. c. Their heat is the same. d. Their heat cannot be compared. e. Other. Please write.....
1.2. Which one of the followings is the reason of your answer in the previous question? A. Temperature depends on the size of the desk. B. Temperature doesn't depend on the size of the desk. C. We can't talk about the temperature of the desks. D. Other. Please write.....	1.2. Which one of the followings is the reason of your answer in the previous question? a. Heat depends on the size of the desk. b. Heat doesn't depend on the size of the desk. c. We can't talk about the heat of the desks. d. Other. Please write.....
1.3. Are you sure about your answers given to the previous two questions? A. Sure. B. Not sure.	1.3. Are you sure about your answers given to the previous two questions? a. Sure. b. Not sure.

<p>2. Answer the following questions for the two same size desk (one is made of iron, the other is wood) resting in a room for a long time.</p>	
<p>2.1. Which one of the desk has higher temperature?</p> <p>A. Wood desk has higher temperature.</p> <p>B. Iron desk has higher temperature.</p> <p>C. Their temperature is the same.</p> <p>D. Their temperature cannot be compared.</p> <p>E. Other. Please write.....</p>	<p>2.1. Which one of the desk has higher heat?</p> <p>a. Wood desk has higher heat.</p> <p>b. Iron desk has higher heat.</p> <p>c. Their heat is the same.</p> <p>d. Their heat cannot be compared.</p> <p>e. Other. Please write.....</p>
<p>2.2. Which one of the followings is the reason of your answer in the previous question?</p> <p>A. Temperature depends on the material of the desk.</p> <p>B. Temperature doesn't depend on the material of the desk.</p> <p>C. We can't talk about the temperature of the desks.</p> <p>D. Other. Please write.....</p>	<p>2.2. Which one of the followings is the reason of your answer in the previous question?</p> <p>a. Heat depends on the material of the desk.</p> <p>b. Heat doesn't depend on the material of the desk.</p> <p>c. We can't talk about the heat of the desks.</p> <p>d. Other. Please write.....</p>
<p>2.3. Are you sure about your answers given to the previous two questions?</p> <p>A. Sure.</p> <p>B. Not sure.</p>	<p>2.3. Are you sure about your answers given to the previous two questions?</p> <p>a. Sure.</p> <p>b. Not sure.</p>

<p>3.1. What is the relationship between heat and temperature in the context above?</p> <p>A. Heat and temperature are the same.</p> <p>B. Temperature is a measure of heat.</p> <p>C. Heat and temperature are not the same. However, they can be used interchangeable in the context above.</p> <p>D. Heat is the energy transferred due to temperature difference.</p> <p>E. Other. Please write.....</p>
<p>3.2. Are you sure about your answer given to the previous question?</p> <p>A. Sure.</p> <p>B. Not sure.</p>

Üç-Aşamalı Isı ve Sıcaklık Testinin Geliştirilmesi ve Uygulanması: Lisans ve Lisansüstü Öğrencileri Örneği

(Özet)

Problem Durumu: Alan yazını, öğrencilerin fizik başarılarına etki eden önemli faktörlerden birinin öğrencilerin o konudaki kavram yanlışları olduğunu göstermektedir. Yeni çalışmalar bu kavram yanlışlarını daha geçerli, güvenilir ve kolay nasıl ölçülebilir üzerinde odaklanmaktadır. Çoğu öğretmen ve araştırmacı başarı testlerinin geliştirilme süreçlerine aşına olmalarına rağmen kavram yanlışsı testinin geliştirilme sürecine yabancıdırlar. Başarı testlerinde sorular rakamsal ve kavramsal içerikliken kavram yanlışsı testlerinde genellikle kavramsaldır. Başarı testlerinde doğru şık seçildiyse puan verilirken kavram yanlışsı testlerinde doğru şık veya kavram yanlışsını ölçen şık seçildiyse puan verilir. Başarı testlerinin geçerliliğini göstermek için soruların hangi hedef davranışları ölçtüğü listelenirken kavram yanlışsı testleri için hangi kavram yanlışları ölçtüğü listelenir. Başarı testleri genelde grup ortalamasına bakarak değerlendirilirken kavram yanlışsı testleri önceden belirlenen ölçütlere göre değerlendirilir. Başarı testlerinin iç güvenirligi genelde kavram yanlışsı testlerinin güvenirligiden büyüktür. Ve son olarak başarı testleri genelde not verme çok az dönüt verme veya sınıflama amaçlı kullanılırken kavram yanlışsı testleri tanılama amaçlı kullanılır. Fizik eğitiminde en tanınmış kavram testi de Kuvvet Kavram Testidir. Bu test çok yaygın olarak alan yazınında kullanılmasına rağmen son zamanlarda gerçekten ne ölçtüğü ile ilgili kritik edilmektedir. Bu testin yanlış sebepli doğru yüzdeleri ve doğru sebepli yanlış yüzdelerinin ne kadar büyük olduğu ile çeldiricilerinin ne kadar çalıştığı sorgulanmaktadır. Aynı zamanda testin ölçmeye çalıştığı alt boyutlarla faktör analizi sonucunda ortaya çıkan alt boyutların örtüşmemesi bu kritikleri arttırmıştır. En tanınmış ve alanda yaygın olarak kullanılan testlerde bile bulunan bu problemleri çözmek için üç-aşamalı testler kullanılmaya başlanılmıştır. Fakat bu testlerin de güvenirlilik katsayıları çok düşük çıkmaktadır.

Araştırmanın Amacı: Bu bağlamda, bu çalışmanın iki temel amacı vardır: 1. Isı ve sıcaklık hakkında kavram yanlışlarını ölçmek için üç-aşamalı bir test geliştirmek ve testin geçerliliğini sağlamak, 2. Bu aracı kullanarak lisans, mastır ve doktora öğrencileri veya mezunlarının yeni bir kodlamayla kavram yanlışlarını ölçmek. Üç-aşamalı test kullanılmasının sebebi, yanlış sebepli doğru ve doğru sebepli yanlış oranlarının rahatça hesaplanabilmesi ve kavram yanlışsının hata ve bilgi eksikliği ile ayırt edilebilmesidir. Yeni kodlamanın kullanılmasının amacı ise bu araçlarda görülen güvenirlilik katsayısı düşüklüğüne ve faktör analizi sonuçlarının ölçülmek istenen boyutlarla örtüşmemesine çözüm bulmaktır.

Araştırmanın Yöntemi: Alan yazınından faydalanarak ısı ve sıcaklık konusu; Türk ve diğer ülke öğrencilerinin öğrenme zorluğu yaşadığı konulardan bir tanesidir. Bu konuda açık uçlu sorular hazırlanarak bir grup öğrenci ile karşılıklı görüşmeler yapılmıştır. Bu görüşmeler sonucunda revize edilen açık uçlu sorular yazılı olarak 36 lise öğrencisinin çözüm ve sebeplerini yazmaları istenmiştir. Daha sonra öğrenci cevapları sınıflandırılarak birinci ve ikinci aşama sorularının seçenekleri oluşturulmuştur. Oluşturulan bu test birkaç fizik eğitim doktora öğrencisi ve iki öğretim üyesi tarafından kontrol edilmiştir. Bu sorularda aynı maddeden yapılmış farklı büyüklükteki masaların ve farklı maddeden yapılmış aynı büyüklükteki masaların hem ısılarını hem de sıcaklıklarını karşılaştırmaları istenmiştir. Her çoktan seçmeli soruya bir boş şık eklenerek öğrencilerin kendi isteklerini de girmeleri sağlanmıştır. Son olarak bu test, yıllardır fizik eğitim sitesi olarak hizmet veren ve günlük ortalama 4000 kişi tarafından ziyaret edilen bir sitede yayımlanmış ve ilgili tartışma forumlarında duyurulmuştur. İki buçuk ayda toplam 3879 kişi testi çevrim içi olarak doldurmuşlardır. Birden fazla defa anketi yollayanlar çıkarıldığında 3405 veri kalmıştır. Bunlardan 1507'si ilköğretim veya lise öğrencisi veya mezunu olduklarını belirtmişlerdir. Bu öğrencilerin verileri bu çalışmada incelenmemiştir. Diğer katılımcıların 1619'u lisans, 219'u mastır ve 60'ı doktora öğrencileri veya mezunları olduklarını belirtmiştir.

Araştırmanın Bulguları: Her soruda verilen boş şık katılımcıların %3,3'ü doldurmuştur. Dolduranların çoğu yazılı olan şıklardaki bilgileri farklı kelimelerle tekrarlamışlardır. Bunlar o şıkları seçmiş olarak kodlanmıştır. Diğerleri ise soruya cevap vermek için gerekli olmayan bazı bilgileri (oda sıcaklığı, masaların ilk sıcaklıkları, masaların kütlelerinin ne oldukları ve ısı kapasiteleri vb.) sormuşlardır. Doğru cevaplar üç-aşamalı sorular ile değerlendirildiğinde aynı cisimden yapılmış farklı büyüklükteki masaların sıcaklıklarını ortalama olarak katılımcıların %63'ü, ısılarını ise %11'i doğru olarak karşılaştırmışlardır. Farklı maddeden yapılmış aynı büyüklükteki masaların sıcaklıklarını ortalama olarak katılımcıların %43'ü, ısılarını ise %10'i doğru olarak karşılaştırmışlardır. Bu yüzdelerle lisans ve mastır öğrencileri arasında önemli bir fark gözlenmez iken bunlar ile doktora öğrencileri arasında doktora öğrencileri lehine bir fark gözlenmektedir. Isı ve sıcaklık arasındaki ilişkiyi ölçen son soruda ise ortalama %55 oranında doğru cevap gelmiştir. Katılımcılar arasında çok farklılık olmamakla birlikte cevaplardaki yanlış sebepli doğru oranı %6, doğru sebepli yanlış oranı %4 ve bilgi eksikliği oranı ise %10 olarak hesaplanmıştır. Soruları normal üç-aşamalı sorular olarak kodladığımızda (her sorunun üç aşamasını bir soru gibi kodlamak) yalnızca sıcaklık ve ısı ile ilgili kavram yanlışlarını ölçerken, bu çalışmaya ait farklı bir şekilde kodladığımızda (sıcaklık ve ısı sorularının her aşamasını bir soru olarak kodlamak) "ısı ve sıcaklık aynıdır" kavram yanlışını ölçebilmektedir. Kavram yanlışlığı

sonuçları incelendiğinde, katılımcılar masanın sıcaklığının daha çok masanın yapıldığı maddeye bağlı olduğunu söylerken, ısısının daha çok kütesine bağlı olduğunu belirtmişlerdir. “ısı ve sıcaklık aynıdır” kavram yanlışları lisans öğrencileri için %5 ile %48 arasında, mastır öğrencileri için öğrencileri için %4 ile %49 arasında ve doktora öğrencileri için %7 ile %40 arasında katılımcıların kavram yanlışısını yalnızca bir durumda veya tüm durumlarda seçmelerine göre değişmektedir. Her iki şekildeki kodlamada da kavram yanlışları oranları tek aşamalı çoktan seçmeli sorular yerine üç-aşamalı çoktan seçmeli sorular ile ölçüldüğünde ortalama olarak %10 azalmaktadır. Bu fark, tek aşamada öğrencilerin hata ve bilgi eksikliklerinden dolayı kavram yanlışısına düşmelerini de kavram yanlışısı olarak ölçmesinden kaynaklanmaktadır. Bu sonuçlar güvenilirlik değerleri ve faktör analizi sonuçları ile desteklenmiştir.

Araştırmanın Sonuçları ve Önerileri: Sonuçlar çoktan seçmeli soruların üç-aşamalı sorulara göre başarı ve kavram yanlışısını ölçmede sınırlılıklarının olduğunu göstermiştir. Başarı ve kavram yanlışları ölçülürken yanlış sebepli doğrular, doğru sebepli yanlışlar ve bilgi eksikliklerine dikkat edilmesi gerektiği ortaya çıkmıştır. Toplam test notu yerine, her kavram yanlışısı için ayrı bir not hesaplanması daha doğru olacaktır. Bu notların güvenilirlik katsayıları ve faktör analizleri ile de desteklenmesi gerekmektedir. Çalışmada geliştirilen test, lisans ve lisansüstü öğrencilerin ısı ve sıcaklıkla ilgili kavram yanlışılarını daha geçerli, güvenilir ve kolay olarak ölçmek için kullanılabilir.

Anahtar Sözcükler: Fizik Eğitimi, üç-aşamalı test, kavram yanlışısı testi, ısı, sıcaklık

Effects of Cooperative Games on Social Skill Levels and Attitudes Toward Physical Education

Osman Gülay*

Dilşad Mirzeoğlu**

Müberra Çelebi***

Suggested Citation:

Gülay, O., Mirzeoğlu, D. & Çelebi, M. (2010). Effects of cooperative games on social skill levels and attitudes toward physical education. *Eurasian Journal of Educational Research*, 40, 77-92.

Abstract

Problem Statement: Physical education (PE) is a part of the general educational program that contributes to the total growth and development of all children. The physical education lessons also contribute to the children's social life and interpersonal relations. Class exercises, games, dances and other activities are generally used for developing children. Many types of games used in PE lessons such as cooperative games. The cooperative games have become popular in recent years since these games emphasize group interaction and positive socialization in a cooperative setting that de-emphasizes competition.

Purpose of Study: The aim of this study was to investigate the effects of 12 week PE class enriched with cooperative games on the ninth grade students' social skills and attitudes toward PE lessons.

Methods: The study involved the ninth grade students in a public elementary school in Bolu

during 2007-2008 educational year. Fortyfour ninth grade female students (22 students in the experimental group and 22 students in the control group) participated in this study. The Social Skills Inventory and The Attitudes towards Physical Education and Sports Scale were administered

This manuscript is a part of unpublished master thesis.

* Ph D, Abant İzzet Baysal University, School of Physical Education and Sport, gulayosman@gmail.com

** Asst. Prof. Dr., Abant İzzet Baysal University, School of Physical Education and Sport, belcesu@yahoo.com

*** Asst. Prof. Dr., Abant İzzet Baysal University, School of Physical Education and Sport, muberracelebi@hotmail.com

to all students. During 12-weeks PE lessons, experimental group participated cooperative games. On the other hand, the control group continued to participate regular PE class based on program with traditional method. The 2x2 (Treatment condition:Experimental/Control x Time:Pre/Post test) repeated measure analysis of variance (ANOVA) and one way analysis of variance (ANOVA) were used to analyze the data.

Findings and Results: Findings indicated significant differences between pre-post tests measurements of subscales of social skills which are emotional expressivity, emotional control, social sensitivity, social control and total social skills. On the other hand, no significant treatment and time x treatment interaction effects were obtained for social skills and attitudes toward PE. In addition, no significant time difference was obtained for attitudes toward PE.

Conclusions and Recommendations: In conclusion, both PE class enriched with cooperative games and physical education courses with traditional methods had similar contribution to the improvement of the students' social skills and attitudes toward PE. In this study, social skills and attitudes of students were only assessed by using quantitative approaches. In future studies, qualitative approach should be used to understand the changes in social skills and attitudes of students.

Keywords: Physical education, social skill, cooperative games, attitude

Physical education is an educational process that contributes to the individual optimal development and well-being by enhancing their skills, fitness, knowledge and attitudes (Bucher & Wuest, 1999). According to Pangrazi (2007, p.5), "physical education is the part of the total educational program that contributes, primarily through movement experiences, to the total growth and development of all children. In that way physical education ensures one's physical, personal and social cohesion". The physical education lessons primarily incorporate the movement education and also contribute to the children's social life and interpersonal relations (Bucher & Koenig, 1983; Gallahue, 1996). Students acquire such with the help of the in-class exercises, games, dances and other activities.

Among these activities, games are the indispensable part of the physical education and enable children to refresh themselves (Boratow, 2006). Though games are considered to be addressing merely the children, they are significant to make students from all ages, gain cognitive, psychomotor and affective skills. Also, games are utilized to design the contents of physical education courses from elementary to secondary school stages of instruction. Especially, cooperative games are distinctive and offer an opportunity to act in a cooperative manner and emphasize on the cooperation, creativity and decision making (Ramsey & Rank, 1997). As reported by Gallahue (1996, p.170) "the cooperative games have become popular in recent years since these games emphasize group interaction and positive socialization in a cooperative setting that de-emphasizes competition". It may be possible to claim that

the cooperative games play an active role in socialization process since these games provides group interaction environment, one-to-one communication, mutual assistance, empathy-building, group action and group awareness (Pangrazi, 2007; Orlick,1979; Orlick, 2006). Kovacs (2008, p.28) also stated that “the cooperative games can bring children, groups and communities together and eliminate fear and the feeling of failure. These games also provide an opportunity for challenge, stimulation and self-validation”.

The previous studies focusing on the effects of games and the cooperative games indicated that children who grow up by playing games are socially more active, are more creative, have richer vocabularies, talk more smoothly (Saracho, 1998), and are happier in doing tasks (Grineski,2002; Orlick, 1981) than those children who grow up without playing games. In addition, cooperative games tend to increase cooperative behaviors of children and help children to behave less aggressively (Orlick, 1976; Orlick, McNally, & O’Hara, 1978; Orlick, 1979; Bay-Hintz Peterson, & Quilitch, 1994; Orlick, 1981). Cooperative games and activities have also been associated with improving peer acceptance, self-esteem (Ames, 1981; Johnson & Johnson, 1985; Madden & Slavin, 1983) and social skills (Pangrazi, 2007). On the other hand, the previous studies (Johnson, Johnson & Scott, 1978; Johnson, Maruyama, Johnson, Nelson & Skon, 1981; Madden & Slavin 1983; Slavin 1990) inconsistently reported improvement with cooperative activities academic performance.

As reported in the above literature review, cooperative games in school and outside life contribute the whole development of children and adolescents. A physical education class in the school is one of the most suitable vehicles for teaching students how to cooperate and compete effectively and enjoyably (Kovacs, 2008; Ramsey & Rank, 1997). Physical education classes through cooperative games also help the socialization of children and adolescents. Hollander, Wood and Hebert (2003) suggested that social skills might be developed by well-prepared health program in schools. Therefore, physical education classes in primary and secondary schools are very important in developing social skills of the students. In order to teach and develop social skills of the students in PE classes, some instructional strategies such as cooperative learning (Dyson, 2001), peer tutoring (Johnson and Ward, 2001), reciprocal style (Mosston and Ashworth, 1996) and social responsibility model (Hellison, 1995) have suggested (cited in Vidoli, 2003).

The previous studies were generally examined the positive contribution of cooperative games to different features of individual such as academic achievement, self-esteem, aggression and peer acceptance but not its possible contribution to social skills and attitudes. Beside previous studies have not examined the effects of cooperative games in P.E. class on the individuals’ affective and psychomotor development. Within our literature search, there is only one study (Ringgenberg, 1998) that examined the effects of cooperative games on classroom cohesion. Ringgenberg (1998) examined the effects of the cooperative games on bowling scores of students and he found that cooperative games improved bowling scores of the individuals, yet, did not affect the amount of cohesion felt among groups. Thus, in the present study, the effects of cooperative games in PE class on the social skills and

attitudes toward PE of secondary school students were studied. The social skill was studied as a variable in this study since good social skills provide students with the ability to make choices that strengthen their interpersonal relationships and lead to success in school. Students with good social skills contribute to a positive and safe school environment. Such skills help students develop resiliency and deal with future crises or other stressful life events (NMHEC, 2002).

In the present study, social skill viewed as a multidimensional concept (Riggio, 1986; 1989) which includes emotional expressivity (EE), emotional sensitivity (ES), emotional control (EC), social expressivity (SE), social sensitivity (SS) and social control (SC). Emotional expressivity relates to skill in sending/encoding nonverbal and emotional messages, but also includes the nonverbal expression of attitudes, dominance, and interpersonal orientation. Emotional Sensitivity is skill in receiving and interpreting the nonverbal messages as well as attentiveness to nonverbal cues. Emotional Control is ability to regulate emotional and nonverbal displays of behavior. Social Expressivity is skill in verbal expression and the ability to engage others in social interaction. Social Sensitivity is the ability to accurately decode/interpret others' verbal communications, as well as knowledge of and sensitivity to the norms governing appropriate social behavior. Social Control is skill in social role-playing and social self-presentation (Perez, Riggio & Kopelowicz, 2007, p.31).

Another study variable in this study is the attitude since attitudes toward physical education has been an important issue in the PE literature recently. Beside, teacher behavior and content of curriculum influenced students' attitude in positive or negative ways to physical education (Figley, 1985; Luke & Sinclair, 1991). The previous studies related to attitudes toward physical education indicated that attitudes scores of male students were higher than female students' scores (Smoll & Schutz, 1980; Şişko & Demirhan, 2002; Kangalgil, Hünük & Demirhan, 2004; Hünük, 2006 and primary school students' attitude scores toward physical education course were higher than secondary school students' scores (Aicenena, 1991; Şişko & Demirhan, 2002). In addition, other studies indicated that sportman students' attitudes scores toward physical education were higher than students who didn't participate in sports (Hünük, 2006; Kangalgil, Hünük & Demirhan, 2004; Alpaslan, 2008). It was assumed that structuring PE class with cooperative games that emphasize cooperation and process of activities could improve the attitudes of students. Baylan (2009) found that there were a statistically significant relationship among social skills, attitudes toward physical education and sport course and self-efficacy level of students in primary and secondary schools. As attitudes of students toward physical education and sport increase, positive and negative social skill levels of students increase, too. Moreover, the result of a research which was done in mentally handicapped children indicated that physical education and sport activities includes cooperate with others, have good social skills and positive attitudes toward physical activities and sports (İlhan, 2008).

In summary, the purpose of this study was to investigate the effects of cooperative games on the ninth grade female students' social skills and attitudes

toward PE classes. It was hypothesized that 12 week cooperative games improved all dimensions of social skills and attitudes of students toward PE class.

Method

Research Design

True experimental design (pretest/posttest equivalent groups design) was used in this study. Participants were randomly and equally assigned to experimental and control groups. Both groups were tested before and after the intervention.

Sample

The population of the study was the ninth grade female students of a secondary public school in Bolu which is the central part of Turkey, in 2007-2008_educational year. The sample was selected for the experimental-control groups among 144 ninth grade female students. The previous studies indicated that female students' attitudes toward PE class were lower than male students' attitudes scores. Therefore, this research was conducted only female students. The Social Skills Inventory and the Attitudes towards Physical Education and Sports Scale were applied to the seven ninth grade classes (144 females) at the beginning of the study. The one-way variance analysis (Anova) were used to determine the differences among all the groups and according to the results of the pre-test scores of attitudes, there were no significant differences on attitudes toward PE among seven classes ($F_{(6-136)} = .910$, $p = .490$). According to this result, it's be said that the calsses had similar attitudes towards physical education lesson. In addition, according to the results of the social skills inventory, emotional sensitivity subscale [$F_{(6-136)} = 2.680$, $p = 0.017$], social expressivity subscale [$F_{(6-136)} = 2.988$, $p = 0.009$] and total social skills [$F_{(6-136)} = 3.185$, $p = 0.006$] showed statistically significant differences among the groups, but no significant differences in the other subscales ($p > 0.05$). The Tukey HSD Test was applied to find the differences in the emotional sensitivity, social expressivity subscales and total social skill level of the female students. According to the result of Tukey HSD test, three groups were eliminated. Then, the experimental and control groups were choosen randomly from the four classes who had similar attitude and social skill scores. Depending on the random selection, the sample of the study were 44 students, 22 female students in the experimental group and 22 female students in the control group.

Research Instruments

The Social Skills Inventory (SSI, Riggio, 1986). The social skills of students' were assessed by using The Social Skills Inventory. It was developed by Riggio in 1986 and was revised in 1989. The inventory has 90 items (15 items per subscales) and each item is responded on 5 point Likert scale. The inventory consists of six subscales with total social skill/competence scales: emotional expressivity, emotional sensitivity, emotional control, social expressivity, social sensitivity, and social control. Total Social Skill (SSI Total) is a general indicator of total social competence that can be

obtained by summing the SSI dimensions. The scores of subscales are ranged from 15 to 75. For total social skill the score is ranged from 90 to 450 (Yüksel, 2004).

The Scale of Attitudes toward Physical Education and Sports (SAPES, Demirhan & Altay, 2001). Attitudes towards Physical Education and Sports Scale (SAPES) developed by Demirhan and Altay (2001) was used in order to assess the attitudes of students towards physical education and sports class. This scale is 5 point likert type scale and it has 24 items in total, 12 of which are affirmative and 12 are negative. The lowest scores in the scale is 24 and highest score is 120. The higher score indicates positive attitudes.

Validity and Reliability

The validity and reliability of The Social Skills Inventory. For the original version of The Social Skills Inventory, the reliability co-efficient of total social scale was .94 and the internal consistency (Cronbach's Alpha) was .89. For SSI subscales the internal consistency were between .63 and .87 (Perez et al, 2007). The Turkish version of SSI has good reliability and validity. The internal reliability of the total social skill scale was .92, and the reliability of the subscales varied from .80 to .89 (Yüksel, 2004).

The validity and reliability of The Scale of Attitudes toward Physical Education and Sports. For The Scale of Attitudes toward Physical Education and Sports, Cronbach Alpha internal consistency coefficient was .93 and the criterion validity coefficient of the scale was .83.

Procedure

After the random selection of the experimental and control groups, experimental group participated two hours cooperative games (80 minutes) per week through 12 week in their regular PE class. The PE class for experimental group started with 20 minute warm up followed by 60 minutes of cooperative games to enhance satisfaction from course, social interaction, and social communication. Some examples of cooperative games are Lion and Cow, Lemonade, Blanket Volleyball, Catch the Dragon's Tail etc.

For example in Blanket Volleyball, two teams use a blanket to toss a beach ball (or other lightweight ball) back and forth over a rope or volleybal net. Every time one team tosses the ball over the net and the other team successfully catches it one collective point is scored. This game is extremely cooperative in structure because every team member is a part of every toss and every catch made by the team. In addition, both teams work together toward a common end, and success is possible only through cooperation (Orlick, 2006). In the other game, Catch the Dragon's Tail players line up with arms around the waist of the person in front. Last one has a handkerchief in her pocket. The player at the head of the line tries to grab the handkerchief. No part of the Dragon may break (Luvmour & Luvmour, 2007,p.25). During the class, teacher emphasize process rather than outcome of games. The teacher also encourages cooperation, sharing, helping behaviors and creativity during the class. The teachers create enjoyable, positive, relax, flexible and likable class atmosphere. The details of cooperative games for 12 week period was given in Table 1.

Table1.*The contents of courses in experimental and control groups*

WEEKS	EXPERIMENTAL GROUP (Name of the Games)	CONTROL GROUP
1. WEEK	Game 1-Hostage Game 2-Cacth Them All Game 3-Pick Up The Balls	Gymnastics, Track and Field (basic gymnastic positions and lining) (sprints-100/200m.)
2. WEEK	Game 1- Big Jump Game 2-Dragon Dodge Ball Game 3-Be Quick	Gymnastics, Track and Field (rhythmical activities) (long jump)
3. WEEK	Game 1-Go Up Game 2-Cacth The Dragon's Tail Game 3-Toby Terrific Turtle	Gymnastics, Track and Field (fundamental skills in floor) (triple jump)
4. WEEK	Game 1-Leon and Cow Game 2-Standing Together Game 3-Lemonade	Gymnastics, Track and Field (apparatus and non-apparatus skills) (shot put)
5. WEEK	Game 1-Base Ball Pass Game 2-Catch The Little Fishes Game 3-Game Of Team Balance	Gymnastics, Track and Field (apparatus and non-apparatus skills) (javelin throw)
6. WEEK	Game 1-Blinds Can Hear And Deafs Can See Game 2-Spider Web Game 3-Snake In The Grass	Track and Field (sprints tournament)
7. WEEK	Game 1-Pass Other Side Game 2-Shape Tag Game 3-Big Pips	Volleyball (overhead pass)
8. WEEK	Game 1-4 Hand And 4 Arm But One Body Game 2-Use That Rope Game 3- Blanket Volleyball	Volleyball (forearm pass)
9. WEEK	Game 1-Still Photograph Game 2-Whell Game 3-Kepp Stand	Volleyball (serving)
10. WEEK	Game 1-Stand Up If U Can Game 2-Destroy It Game 3-Nine Rocks	Volleyball (match)
11. WEEK	Game 1-Pass Inside Ring Game 2-Tag With Baloon Game 3-Don't Step The Ground	Measurement and Evaluation (volleyball skills)
12. WEEK	Game 1-We Are Buddy Game 2-Watch Out Game 3-Tarzan	Matches (volleyball, basketball and handball)

The control group participated 80 minutes regular PE class throughout 12 week period. PE class was structured in accordance with the annual curriculum prepared by the physical education teacher. The traditional PE class was teacher centered, generally teacher used command style, demonstration, lecturing and discussion about subjects of the curriculum during the class (Table 1). For both experimental group and control group PE class was taught by same teacher who is one of the researchers of the present study.

Data Analyses

Data that were obtained from both experimental and control groups was analyzed by using 2 x 2 (Treatment: Experimental/Control x Time:Pre/Post test) Repeated Measure Analysis of Variance (ANOVA). Besides, the one-way analysis of variance (ANOVA) and the Tukey HSD test were used to determine the experimental and control groups at the beginning of the study. The level of significance was accepted as 0.05.

Findings and Results

The results of a 2 x 2 (Treatment Condition: Experimental/Control x Time: Pre/Post test) repeated measure analysis of variance test (Anova) indicated significant differences in emotional expressivity ($F_{1,42}=5.94$, $p<.05$; $n^2=0.13$, power=0.66); emotional control ($F_{1,42}=4.33$, $p<.05$; $n^2=0.94$, power=0.530); social sensitivity ($F_{1,42}=17.08$, $p<.01$; $n^2=0.29$, power=0.981); social control ($F_{1,42}=7.44$, $p<.01$; $n^2=0.15$, power=0.760) between pre-test and post-test measurements (Table 2). In addition, there was a significant time difference in total social skill scores ($F_{1,42}=5.77$, $p<.05$; $n^2=0.12$, power=0.651). The significant time differences between pre-post tests has indicated that the emotional expressivity, social sensitivity, social expressivity and total social skills of all students both in experimental and control increased while emotional control scores decreased through 12 weeks period . ANOVA also showed no significant time differences in the attitudes toward PE and sports ($p >.05$).

2 (Treatment Conditions) x 2 (Time) repeated measure ANOVA also showed no significant time differences in the attitudes toward PE and sports ($p>0.05$). In addition, Treatment condition x Time interactions was not significant for social skill subscales and attitudes toward PE and sports ($p>.05$). Nonsignificant interaction effects revealed that changes in dimensions of social skills and attitudes toward PE and sport scores over 12 weeks did not differ with regard to type of treatment that received.

Table 2.

Means and Standart Deviations for Social Skills and Attitudes Based on the Time of Measurements

Subscales	Groups	Pre-test		Post-test		Time Effect F	p
		M	Sd	M	Sd		
<i>Social Skills</i>							
Emotional Expressivity	Experimental	39.27	8.50	44.41	7.16	5.94	.019
	Control	41.18	7.81	43.68	6.51		
Emotional Sensitivity	Experimental	47.82	9.13	51.36	10.14	2.63	.112
	Control	49.41	9.96	50.27	8.40		
Emotional Control	Experimental	46.41	10.31	43.77	7.83	4.33	.043
	Control	46.14	11.14	41.50	6.56		
Social Expressivity	Experimental	44.18	8.61	47.23	7.37	1.36	.250
	Control	44.59	8.57	46.73	8.89		
Social Sensitivity	Experimental	45.14	7.95	50.45	7.92	17.08	.000
	Control	47.27	8.98	53.82	7.96		
Social Control	Experimental	43.91	9.88	46.86	8.06	7.44	.009
	Control	40.27	8.26	48.18	8.55		
Total Social Skills	Experimental	263.50	43.13	285.41	28.22	5.77	.021
	Control	274.09	47.81	282.86	23.37		
<i>Attitudes toward PE and sports</i>							
	Experimental	94.91	14.53	95.82	14.98	.692	.410
	Control	91.05	14.76	93.82	14.98		

Discussion

The primary goal of this study was to test the possible effects of cooperative games in PE class on the social skills and attitudes toward PE and sport in secondary school female students.

Analysis indicated that emotional expressivity, social sensitivity, and total social skill scores of students in both cooperative game PE class and traditional PE class improved from pre to posttest. However, no changes obtained in emotional sensitivity and social expressivity from pre to post test. The significant time differences indicated that the emotional expressivity, social sensitivity and total social skills of all students (both in experimental and control) increased while emotional control and social control scores decreased through 12 weeks period. In other words, 12 weeks PE class (annual program with traditional vs cooperative games) participation improved interpersonal and communication skills of students.

Generally, these findings could be supported by the studies that reported positive effects of PE on socialization process of children and adolescents (İlhan, 2008). This might also be explained by the contribution of the physical education lessons to the students's social life and interpersonal relations (Serbes, Yüret, & Topkaya, 2004). A physical education class in the school is one of the most suitable vehicles for teaching students how to cooperate and compete effectively and enjoyably (Kovacs, 2008; Ramsey & Rank, 1997). Furthermore, in generally the nature of PE class involves games, physical activities, fun, and cooperation. That's why studies (Saracho, 1998; Bay-Hintz et al, 1994; Orlick et al, 1978; Orlick, 1981; Koçkar, Girmen, Anılan, & Öztürk, 2004; Avcıoğlu, 2001) that reported positive effects of games and cooperative games on the development of students could support the present study. For example, Hall (1987) states that group games play a vital role in the children's ability to develop strong communicative skills (cited in Mangır & Aktaş, 1993). Above explanations could also apply to the obtained time changes in total social skill scale in the present study.

Specifically, time changes in emotional expressivity indicated that secondary school female students improved their skill in sending/encoding nonverbal and emotional messages learn to express their attitudes, dominance, and interpersonal orientation through 12 week PE class participation. The change in emotional expressivity is expected since the nature of PE class provides a chance to students to relax and express themselves easily. On the other hand, findings indicated reduction in emotional control score of female students from pre to post test. If an individual has a high level of emotional control, s/he is observed to control him/herself very well and his/her self-respect to be high accordingly. A person with a high level of emotional control is able to pretend to be so well pleased, conceal the actual feelings even not feeling so indeed, and also when joked, s/he is able to show whatever mimes and gestures are needed to conceal. Whereas in physical education classes, as a course requirement, pupils get the opportunity for expressing their emotions, either positive or negative, with no question or hesitation. Therefore, this outcome ought to be quite natural, as they are given the credit for being able to live what they indeed feel in physical education classes. In this way, the regression out of the post-test of emotional control subscale is to be expected.

For social subscales of social skills, findings indicated significant improvement in social sensitivity and social control. This means that through PE class, students learn how to become more respectful towards one another, more aware of how other people feel (Orlick, 1979). Physical education lessons are likely to be conducive in a way to enable students to develop feeling, value, attitude and social behavior in an affective manner (Rink, 1998). These results may have occurred due to the fact that students are provided during physical education with opportunities to act in unison and be engaged in team-works. These findings were also supported by the study of Kalkan and Sardoğan (2003).

Analysis also revealed no significant time, treatment and treatment of time interaction effects on the attitudes toward PE and sports. This analysis indicated that attitudes toward PE did not change over 12 weeks and also changes in attitudes

scores over time did not change whether or not participating cooperative game PE class or annual program with traditional PE class. In other words, the cooperative games and the annual program with traditional method did not have a prominent effect on the attitudes of the participants. This might be explained by high pre-test attitudes scores of participants. The pre-test means indicated the possible ceiling effect that all groups had higher scores which approach to highest scores of scale. But, in this research, all participants were female. According to the results of some studies, the attitude scores of female students toward PE course are lower than male students' scores (Smoll & Schutz, 1980; Şişko & Demirhan, 2002; Kangalgil, Hünük & Demirhan, 2004; Hünük, 2006). Another reason might be stable nature of attitude. As Tavşancıl (2002) mentioned, attitudes tend to resist to any changes despite peoples' continuous efforts to change each other's attitudes. Eventhough at a slow pace, attitudes, nevertheless, may undergo changes in spite of this resistance. Considering these explanations, it may be true to assert that the duration of the study may have been insufficient for students' attitudes to experience any changes.

Contrary to the hypothesis of this study, improvement in the social skills and attitudes toward PE over 12 weeks did not differ with regards to type of PE class. The results failed to find significant effect of cooperative games on the social skill and attitudes toward PE. In other words, the physical education enriched with the cooperative games and the physical education restricted to the annual program with traditional method, when assessed in terms of social skill development, had a similar effect. La Greca (1993) defined social skills as positive social behaviors contributing to start-up and continuance of positive social interactions (cited in Ekinci, 2006). Social skills play a key role in individual's relations with his/her environment and the society as well. Because the students find more opportunities in the course of physical education to play, share, to take part in mutual assistance, act as leader, obey the leader, develop communication and so on, it may be the case that the social skill levels of the experimental and control groups did not exhibit a significant difference.

Conclusion

In conclusion, physical education enriched with the cooperative games and physical education lesson, which was taught within the limits of the present curriculum and with the traditional, had similar contribution to the improvement of the students' social skills and attitudes toward PE and sports. On the other hand, generally PE classes have remarkable contribution to the improvement of social skills. These results should be carefully interpreted by considering the limitations of the present study. First limitation of the present study is the time of intervention. It may be said that 12 week is not enough to see possible effects of cooperative games on social skills and attitudes. In future studies longer time period should be used to test the time effects of cooperative games. Second, possible gender effects did not consider in this study. The gender might be confounding variable. This should be considered in the future studies. Third, nonsignificant results could be attributed to limited sample size. It's obvious that sample size has great influence on the statistical significance of results.... (Gravetter & Wallnau, 2004). The larger sample size should

be used in further studies. In this study, social skill and attitudes of students were only assessed by using quantitative approach. In future studies, qualitative approach should be used to understand the changes in social skills and attitudes of students.

References

- Aicinena, S.J. (1991). The teacher and student attitudes toward physical education. *The Physical Educator*, 48, 28-32.
- Alpaslan, S. (2008). *Ortaöğretim öğrencilerinin beden eğitimi öğretmenlerinin sergilediği öğretim davranışlarına ilişkin algıları ve öğrencilerin bu derse karşı geliştirdikleri tutumlar*. [The perception of high school student's toward physical education teachers's behavior and attitudes of students toward physical education course]. Unpublished master dissertation, Abant İzzet Baysal Üniversitesi, Sosyal Bilimler Enstitüsü, Bolu.
- Avcıoğlu, H. (2001). *İşitme engelli çocuklara sosyal becerilerin öğretilmesinde işbirlikçi öğrenme yaklaşımı ile sunulan öğretim programının etkililiğinin incelenmesi* [The effect of physical education upon the socialization levels of mentally handicapped children]. Unpublished doctoral dissertation, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Özel Eğitim Anabilim Dalı, Ankara.
- Ames, C. (1981). Competitive versus cooperative reward structures: The influence of individual and group performance factors on achievement attributions and affect. *American Educational Research Journal*, 18, 273-287.
- Bay-Hintz A. K., Peterson R. F., & Quilitch H. R.(1994). Cooperative games: a way to modify aggressive and cooperative behaviors in young children. *Journal of Applied Behavior Analysis*, 17 (3), 435-446.
- Baylan, M. (2009). *İlköğretim 2. kademe ve orta öğretim kurumlarındaki öğrencilerin beden eğitimi dersine yönelik tutumları, sosyal beceri ve öz yeterlik düzeylerinin karşılaştırılması*. [The comparison of primary and secondary school students' attitudes toward physical education course, social skills and self-efficacy]. Unpublished doctoral dissertation, Ege Üniversitesi, Sağlık Bilimleri Enstitüsü, İzmir.
- Bucher, C. A. & Wuest, D. A. (1999). *Foundations of physical education and sport*. The McGraw-Hill Companies, New York.
- Bucher, C. A. & Koenig, C. R. (1983). *Method and material for secondary school physical education*, The C.V. Mossey Company, Saint Louis.
- Boratow, P. R. (2006). *100 Soruda Türk Folkloru* [Turkish folklore in 100 questions]. K Kitaplığı Yayınevi, İstanbul.
- Demirhan, G., & Altay, F. (2001). Lise birinci sınıf öğrencilerinin beden eğitimi ve spora ilişkin tutum ölçeği II [The scale of attitudes of ninth grade students toward physical education and sports]. *Hacettepe Üniversitesi Spor Bilimleri Dergisi*, 12 (2), 9-20.

- Ekinci, Y. (2006). *İlköğretim Okulu Yöneticilerinin Sosyal Beceri Düzeylerine Göre Öğretmenlerin İş Doyumu ve İş Stresinin Karşılaştırılması* [Relationship between job satisfaction and organizational stressors of principals in elementary schools]. Unpublished master dissertation, Gaziantep Üniversitesi, Sosyal Bilimler Enstitüsü, Eğitim Bilimleri Ana Bilim Dalı, Gaziantep.
- Figley, G. E. (1985). Determinants of attitudes towards physical education. *Journal of Teaching in Physical Education*, 4, 229-240.
- Gallahue, D. L. (1996). *Developmental physical education for today's children*, Brown&Benchmarks Publishers, Third Edition, Boston.
- Gravetter, F.J. & Wallnau, L.B. (2004). *Statistics for the behavioral sciences*. USA: Thomson, Wadsworth.
- Grineski, S. (2002). *Beden eğitiminde işbirliğiyle öğrenme* [Cooperative learning in physical education]. (Trans: Selim Yeniçeri). İstanbul: Beyaz Yayınları.
- Hollander, D.B., Wood. R.J., & Hebert, E.P. (2003). Protecting students against substance abuse behaviors: integrating personal and social skills into physical education. *The Journal of Physical Education, Recreation and Dance*, 74:5, 45-48.
- Hünük, D. (2006). *Ankara ili merkez ilçelerindeki ilköğretim ikinci kademe öğrencilerinin beden eğitimi dersine ilişkin tutumlarının sınıf düzeyi, öğrenci cinsiyeti, öğretmen cinsiyeti ve spora aktif katılımları açısından karşılaştırılması*. [Comparison of middle school students attitudes toward physical education with regard to grade, student's gender, teacher's gender and sport participation]. Unpublished master dissertation, Hacettepe Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara.
- İlhan, L. (2008). *Eğitilebilir zihinsel engelli çocuklarda beden eğitimi ve sporun sosyalleşme düzeylerine etkisi* [The effect of physical education upon the socialization levels of mentally handicapped children]. *Kastamonu Eğitim Dergisi*, 16:1, 315-324.
- Johnson, D. W., Johnson, R. T., & Scott, L. (1978). The effects of cooperative and individualized instruction in students attitudes and achievement. *Journal of Social Psychology*, 104, 207-216.
- Johnson, D. W., Maruyama, G., Johnson, R. T., Nelson, D., & Skon, I. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analyses. *Psychological Bulletin*, 89, 47-62.
- Johnson, D.W. & Johnson, R.T. (1985). The internal dynamics of cooperative learning groups. In R. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb and R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 103-124). New York: Plenum.
- Kalkan, M. & Sardoğan, M. E. (2003). İnsan ilişkileri becerilerine dayalı bir eğitim programının grup üyelerinin sosyal beceri düzeylerine etkisi [The effect of a training program which is based on human relation skills through the level of the social skills of the group members]. *Psikiyatri Psikoloji Psikiyatrik Psikofarmakoloji (3P) Dergisi*, 11, (2) 137-142.

- Kangalgil, M., Hünük, D., & Demirhan, G. (2004). Spor yapan ve yapmayan kız ve erkek öğrencilerin beden eğitimi ve spora ilişkin tutumları [Attitude of students, who are engaged in sports and who are not, towards physical education and sports]. *The 10th ICHPER-SD Europe Congress & The TSSA 8th International Sports Science Congress*, November 17-20, Antalya, 137.
- Koçkar, M., T., Girmen, P., Anılan, H., & Öztürk, A. (2004). *Geleneksel çocuk oyunlarının eğitimsel değeri [Educational value of traditional children's games]*. Retrieved April 8 2008 from <http://www2.ogu.edu.tr/~mtkockar/cocukoyunlari.htm>.
- Kovacs, K. (2008). Cooperative games in education, The 10th International Sports Science Congress, *Proceedings-I*. October 23-25, Bolu, Turkey, pp:28-29.
- Luke, M., & Sinclair G. (1991). Gender difference in adolescent attitudes toward school physical education. *Journal of Teaching in Physical Education*, 11, 31-46.
- Luvmour, J., & Luvmour, B. (2007). *Everyone wins cooperative games and activities*. New Society Publishers.
- Madden, N, A. & Slavin, R, E. (1983). Mainscreaming students with mild handicaps: Academic and social outcomes. *Reviews of Educational Research*, 53, 519-569.
- Mangır, M. & Aktaş, Y. (1993). Çocuğun Gelişiminde Oyunun Önemi, *Yasadıkça Eğitim Dergisi*, 26, 14-18.
- National Mental Health & Education Center (NMHEC) (2002). *Social skills: promoting positive behavior, academic success and school safety*. Retrieved June, 7, 2004., from http://www.naspcenter.org/fastsheets/socialskills_fs.html
- Orlick, T. D. (1976). Games of acceptance and psychosocial adjustment. In D. Smith & M. Bar-Eli (Eds), *Essential readings in sport and exercise psychology* (pp.126-134).
- Orlick, T. D. (1981). Positive socialization via cooperative games. *Developmental Psychology*, 17 (4), 426-429.
- Orlick, T. D., McNally, J., & O'Hara, T. (1978). Cooperative games: systematic analysis and cooperative impact. In D. Smith & M. Bar-Eli (Eds), *Essential Readings in Sport and Exercise Psychology* (pp.135-154).
- Orlick, T. (1979). Childrens's games: Following the path that has heart. *The Elementary School Guidance and Counselling Journal*, 14 (2), 156-161.
- Orlick, T. D. (2006). *Cooperative games and sports: Joyful activities for everyone*. Human Kinetics, USA: Champaign.
- Pangrazi, R. P. (2007). *Dynamic physical education for elementary school children*. San Francisco: Pearson Education In.
- Perez, J.E., Riggio, R.E., & Kopelowicz, A. (2007). Social skill imbalances in mood disorders and schizophrenia, *Personality and Individual Differences*, 42, 27-36.
- Ramsey, G. & Rank, B. (1997). *Rethinking youth sports*. Retrieved June 11 2009 from http://findarticles.com/p/articles/mi_m1145/is_n12_v32/ai_20127496/

- Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology*, 51, 649-660.
- Riggio, R. E. (1989). *Manual for the social skills inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Ringgenberg, S. W. (1998). *The effects of cooperative games on classroom cohesion*. Published master of science thesis, University of Wisconsin, Lacrosse.
- Rink, E. J. (1998). *Teaching physical education for learning*, Mc-Graw Hill Company, Boston.
- Saracho, O. N. (1998). Socialization factors in the cognitive style and play of young children, *International Journal of Education and Research*. 3 (29), 263-276.
- Serbes, Ş., Yüret, E., & Topkaya, İ. (2004). İlk ve orta öğretim beden eğitimi öğretimi programının eğitim amaçları açısından, genel amaç ve amaç (hedef) içerikleri analizi [The content analysis of learning objectives pertaining to primary and secondary school physical education teaching program] XIII. *National Educational Sciences Conference*, İnönü Üniversitesi, Eğitim Fakültesi, Malatya,
- Slavin, R. E. (1990). *Cooperative learning: theory research and practice.*, NJ: Prentice-Hall, Englewood Cliffs.
- Smoll, F. L. & Schutz, R. V. (1980). Children's attitudes towards physical activity: longitudinal analysis. *Journal of Sport Psychology*. 2, 137-147.
- Şişko, M. & Demirhan, G. (2002). İlköğretim okulları ve liselerde öğrenim gören kız ve erkek öğrencilerin beden eğitimi ve spor dersine ilişkin tutumları. [The attitudes of male and female students from the primary and secondary schools towards physical education courses]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 23, 205-210.
- Tavşancıl, E. (2002). Tutumların ölçülmesi ve spss ile veri analizi [Attitudes measurement and data analysis with SPSS]. *Nobel Yayın Dağıtım*, Ankara.
- Vidoli, C. (2003). Teaching your class social skills: an explicit strategy. *Teaching Elementary Physical Education*, November,20-22.
- Yüksel, G. (2004). Sosyal beceri envanteri el kitabı [Social skills inventory manual]. *Ankara: Asil Yayın Dağıtım*.

İşbirlikli Oyunların Öğrencilerin Sosyal Beceri Düzeylerine ve Beden Eğitimi Dersi Tutumlarına Etkisi

(Özet)

Problem Durumu: Beden eğitimi çocukların büyüme ve gelişimlerinde öncelikli olarak hareket deneyimlerine katkıda bulunan bütün eğitsel programın bir parçasıdır. Beden eğitimi dersleri öncelikli olarak hareket eğitimini içerir ve öğrencilerin sosyal yaşamlarına ve kişiler arası ilişkilerine katkıda bulunur. Beden eğitimi derslerinde sınıf içerisinde uygulanan oyunlar, danslar ve diğer aktivitelerle öğrencilere bu özellikler kazandırılır. Beden eğitimi derslerinde uygulanan oyunlar çok çeşitlidir. Örneğin, işbirlikli oyunlar son zamanlarda oldukça popüler olmuştur çünkü işbirlikli oyunlar bireysel olarak yarışmayı desteklemeyen işbirlikli bir kurulum içinde grup etkileşimini ve olumlu sosyalleşmeyi içermektedir.

Araştırmanın Amacı: Bu araştırma beden eğitimi derslerinde 12 hafta boyunca oynatılan işbirlikli oyunların 9. sınıf öğrencilerinin sosyal becerilerine ve beden eğitimi dersine karşı tutumlarına etkisini incelemek amacıyla yapılmıştır.

Araştırmanın Yöntemi: Araştırma 2007-2008 öğretim yılında Bolu ili Merkez ilçede bulunan bir ortaöğretim kurumunda öğrenim gören 9. sınıf kız öğrencileri üzerinde yürütülmüştür. Araştırmaya 44 öğrenci (deney grubunda 22 öğrenci ve kontrol grubunda 22 öğrenci) katılmıştır. Araştırmada veri toplama araçları olarak Sosyal Beceri Envanteri (SBE) ile Beden Eğitimi ve Sporda Tutum Ölçeği (BESTÖ) kullanılmıştır. 12 haftalık süreçteki beden eğitimi derslerinde deney grubu öğrencilerine işbirlikli oyunlar oynatılmış, kontrol grubuna ise yıllık öğretim programı doğrultusunda geleneksel yöntem ile beden eğitimi dersi işlenmiştir. Araştırmada elde edilen verileri çözümlemek için 2x2 (İşlem:Deney/kontrol x Süre:Ön/son test) tekrarlı ölçümler için varyans analizi (ANOVA) ve tek yönlü varyans analizi (ANOVA) istatistiksel teknikleri kullanılmıştır.

Araştırmanın Bulguları: Araştırma sonuçlarına göre, sadece sosyal becerinin alt boyutlarında (duyuşsal anlatımcılık, duyuşsal kontrol, sosyal kontrol, sosyal duyarlılık ve toplam sosyal beceri) öntest-sontest puanları arasında anlamlı fark bulunmuştur. Ancak, öğrencilerin sosyal beceri ve beden eğitimi ve spora karşı tutum puanlarında işlem ve işlem x süre ilişkisinin etkisi önemli düzeyde bulunmamıştır. Bununla birlikte, öğrencilerin beden eğitimi ve spora karşı tutumlarında da süre göz önüne alındığında anlamlı fark bulunmamıştır.

Araştırmanın Sonuçları ve Önerileri: Sonuç olarak, hem işbirlikli oyunlarla, hem de geleneksel anlayışla işlenen beden eğitimi derslerinin öğrencilerin sosyal beceri düzeylerini benzer düzeyde geliştirdiği söylenebilir. Bu çalışmada öğrencilerin sosyal becerileri ve beden eğitimi dersine karşı tutumları sadece niceliksel yöntemlerle ölçülmüştür. Bundan sonraki çalışmalarda, öğrencilerin sosyal becerilerindeki değişiklikler ve tutum düzeylerinin niteliksel yöntemlerle ölçülmesi önerilebilir.

Anahtar Kelimeler: Beden eğitimi, sosyal beceri, işbirlikli oyunlar, tutum

Students' Affective Characteristics and Their Relation to Mathematical Literacy Measures in the Programme for International Student Assessment (PISA) 2003

Çiğdem İş Güzel*

Giray Berberoğlu**

Suggested Citation:

İş Güzel, Ç., & Berberoğlu, G. (2010). Students' affective characteristics and their relation to mathematical literacy measures in the Programme for International Student Assessment (PISA) 2003. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 40, 93-112.

Abstract

Background: This study evaluates students' affective factors that are related to the mathematical literacy skills assessed by the Programme for International Student Assessment (PISA) 2003.

Purpose: The purpose of the study is to test a linear structural model to investigate affective variables that are related to the mathematical literacy skills of 15-year-old Turkish students in the PISA 2003.

Design and Methods: The PISA data set was analyzed for Turkish students within the framework of linear structural modeling. The affective variables that are presumably related to mathematical literacy skills were assessed by the student questionnaire in the PISA 2003 and were considered in the proposed model. The following were used as variables to explain the mathematical literacy measures of students: *Interest in and Enjoyment of Mathematics, Instrumental Motivation in Mathematics, Anxiety in Mathematics, Self-Efficacy in Mathematics, Self-Concept in Mathematics, Sense of Belonging at School, and Disciplinary Climate in Mathematics Lessons*. Additionally, the impacts of *Sense of Belonging at School* and *Disciplinary Climate in*

* Dr., Cito Türkiye Office, Middle East Technical University Metropolis, Galyum Blok, B13, 06531, Middle East Technical University, Ankara, TURKEY, cigdem.isguzel@gmail.com

** Prof. Dr., Secondary Science and Mathematics Education, Faculty of Education, Middle East Technical University, 06531, Ankara, TURKEY, giray@metu.edu.tr

Mathematics Lessons on Mathematical Literacy were also tested in the proposed model. Thus, both *Sense of Belonging at School* and *Disciplinary Climate in Mathematics Lessons* were treated as exogenous and endogenous variables.

Results: The greatest relationship was found between *Self-Efficacy in Mathematics* and *Mathematical Literacy*. Other significant relationships with *Mathematical Literacy* were found with the latent variables *Interest in and Enjoyment of Mathematics*, *Anxiety in Mathematics*, and *Disciplinary Climate for Mathematics Lessons*. However, students who indicated positive feelings about interest in and enjoyment of mathematics performed lower than students who reported less interest in and enjoyment of mathematics on the mathematical literacy scale. On the other hand, *Interest in and Enjoyment of Mathematics* seemingly had a small but positive effect on *Mathematical Literacy* measures through the *Disciplinary Climate in Mathematics Lessons* latent variable.

Conclusion: In the present study, a total variance of 42% suggests that the mathematical literacy measure can be used to indicate the importance of affective variables in explaining the academic performance of students. Turkish students have positive attitudes towards mathematics with less confidence and higher anxiety levels when compared to students from other participating countries. Somehow, their positive attitude was not connected with a better academic performance in the educational system. Evidence suggests that the classroom climate is negatively influenced by the high anxiety and low confidence levels of the students. This might also cause classroom management problems for teachers dealing with students with differing academic performance levels and backgrounds in mathematics.

Keywords: Programme for International Student Assessment (PISA); Mathematical literacy; Confirmatory factor analysis; Structural equation modeling; Turkey

Influential factors that affect the achievement level of students in different subject matter areas have been extensively studied in the field of educational research. Among all of the different subject matter areas, mathematics has drawn the attention of many researchers since a majority of students in some countries perform at levels far below that of students from other participating countries in international comparative studies, such as the Programme for International Student Assessment (PISA). For instance, almost 28% of the students in Turkey could not even achieve a score consistent with the minimum proficiency level as defined for the PISA scale (National Education Publications, 2005). The national mean for Turkish students is 423, which is almost one standard deviation below the Organisation for Economic Co-operation and Development (OECD) mean (OECD Publications, 2004a).

In an effort to understand the reasons behind low achievement in mathematics, researchers have focused on student related factors. Among many research studies on this subject, the following topics were extensively studied as predictors of mathematics achievement: academic and mathematics self-concepts (Marsh, 1986); mathematical ability perceptions, performance expectancies, value perceptions, and mathematics anxiety (Eccles et al., 1990); home environment, motivation, mathematics attitude, and classroom environment (Reynolds & Walberg, 1992); attitude towards mathematics on the basis of feelings of enjoyment, feelings of difficulty, and perceptions of mathematics (Ma, 1997); critical events, inattentiveness, student attitudes based on enjoyment of school, teacher responsiveness, and usefulness (Hill & Rowe, 1998); family context, learning experiences, self-efficacy, and interest (Ferry et al., 2000); and antecedent variables, perceived importance of mathematics, and attitudes towards mathematics (Abu-Hilal, 2000).

In addition to student related characteristics, teacher and school related characteristics and their impacts on mathematics achievement have also been considered by some researchers (Bos & Kuiper, 1999; İş Güzel & Berberoğlu, 2005; C. Papanastasiou, 2000, 2002; E. Papanastasiou, 2002; Schreiber, 2002). These studies clearly emphasized the importance of the affective characteristics of students to explain achievement in mathematics. For instance, as students become more interested and motivated and develop self efficacy in learning mathematics, they might be developing more positive feelings about school, which fosters more disciplinary climate in the classroom (Dorman, 2001; Fraser, 1986). Through interaction with mathematics related activities, students will likely develop greater interest and motivation in learning mathematics at an early age. As students get older this type of predisposition could shape their overall feelings about school, thereby perhaps creating a more peaceful environment in the classroom. As a result, a peaceful and disciplined climate in the classroom is expected to improve learning. Thus, within this framework, an investigation of the relationships between affective variables and school related attitudes and classroom climate and their overall impact on mathematical performance of the students is worthwhile.

The PISA provides information beyond students' learning achievement in mathematics. Specifically, the PISA provides further opportunities to understand mathematics related performances within the structure of the literacy concept. Understanding the literacy concept and the mathematical literacy skills of students with respect to various student affective characteristics is seemingly quite important as a means to enhance school curricula to include the content and activities that are crucial to foster necessary life skills that require higher order cognitive processes. These skills are not completely independent of school curricula; rather, they build on basic skills covered in mathematics classes and are used within the context of daily life situations.

Thus, the present study is intended to test a linear structural model to investigate student related affective variables that are associated with school related factors and mathematical literacy skills for Turkish students that were included in the PISA 2003.

The findings will presumably increase our understanding of the low performance level of the Turkish students within the PISA 2003.

Method

Sample

An age-based definition for the target population of the PISA was used; specifically, all students were between 15 years 3 months and 16 years 2 months of age. The mean age of students from all OECD countries was 15 years and 8 months (OECD Publications, 2004a, p.320), and the mean age of Turkish students was 15 years 9 months. A two-stage stratified sample design was used for the PISA (OECD Publications, 2005). The first stage consisted of selecting the individual schools for 15-year-old students, and the second consisted of selecting the students within the sampled schools (OECD Publications, 2005). As a result of this selection process in Turkey, 4,855 students participated in the sample group from a population of 719,702 students in the desired target population within the nation. The sample included 2,090 female and 2,765 male students. The grade levels of the Turkish students ranged from 7th to 12th grade.

Instruments

Responses by the Turkish students on the two PISA instruments, namely, the Student Questionnaire and the Mathematical Literacy Assessment, were used in the present study. The Student Questionnaire contains questions related to information about the students and their family backgrounds (OECD Publications, 2003, 2004a). The mathematical literacy of the students was assessed through open-constructed response items, closed-constructed response items, short response items, complex multiple-choice items, and multiple-choice items on the PISA project (OECD Publications, 2004a). PISA provides estimates of five plausible values for the mathematical literacy score. In this study, all of the five overall mathematical literacy plausible values were used as observed variables to represent the *Mathematical Literacy* latent variable in the path analytic model. The other latent variables came from the questionnaire items used in the Student Questionnaire.

Data Analyses

The data files that were used in this study were downloaded from the PISA International Database on the PISA Web Site. The statistical analyses were conducted using the following steps: (1) Groups of items were selected to form the latent variables for the path analytic model in line with the PISA framework from the Student Questionnaire; (2) Latent variables were evaluated through the confirmatory factor analyses; (3) Covariance matrices were constructed with the observed variables; (4) The fit of the path analytic model with latent variables was tested.

In the present study, LISREL 8.30 for Windows (SSI Inc., 1999) with the SIMPLIS command language was used to analyze the data. The maximum likelihood estimation method was used for all of the LISREL analyses. In the LISREL analyses, all of the covariance matrices were based on the pairwise deletion of the missing cases. This approach, i.e., pairwise deletion, was used to keep as many cases as possible in the data

file. This might affect the representativeness of the sample, but the missing cases do not exceed 7 percent on average of the data file. Thus, this low amount of missing cases should not jeopardize the sampling adequacy in the analyses. To identify influential points, Cook's Distance values were calculated and inspected, and no influential points were observed. For the assessment of the fit of the model to the data, the Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI), Standardized Root Mean Square Residual (S-RMR), and Root-Mean-Square Error of Approximation (RMSEA) were used to interpret the fit of the model proposed (Browne & Cudeck, 1993; Gerbing & Anderson, 1993; Jöreskog & Sörbom, 1993; Kelloway, 1998; Schumacker & Lomax, 1996). Good model fit interpretations are possible if the GFI and AGFI index values are above .90 and S-RMR and RMSEA index values are below .05. However, values of 0.08 or less for the RMSEA and S-RMR indicate a reasonable error of approximation (Browne & Cudeck, 1993).

Dimensions of the Student Questionnaire Items in the Turkish Data Set

In order to establish the latent variables for the LISREL model, the questionnaire items were grouped conceptually in line with the PISA framework. In this particular study, the researchers preferred to develop the sub-dimensions by re-analyzing the questionnaire items for the Turkish sample even though PISA 2003 provides indexed values for the construct that is assessed by the Student Questionnaire. First, in a path analytic model with latent variables, a group of observed variables should be described for the construct under investigation. This requires questionnaire items that are highly correlated to the latent variable. Second, since the PISA Student Questionnaire is offered in different languages and different cultural settings, an intra-cultural grouping of items was used to avoid possible error coming from translation and culture specific contextual factors. This might restrict the cross cultural generalizations of the findings, but for intra-cultural interpretations of data, the results could be more valid for the Turkish sample specifically.

When grouping the questionnaire items conceptually, the PISA framework was used. In this process, out of the 26 dimensions defined in the PISA framework, seven dimensions that are related to the conceptual model of the present study were considered. For the conceptually grouped items, a confirmatory factor analysis with a seven-factor solution was carried out in order to assess the data-model fit. For the seven-factor model, the GFI, AGFI, S-RMR, and RMSEA were 0.93, 0.91, 0.039, and 0.051, respectively. However, for a better fit, some covariance terms were added into the model by considering the modification indexes (Kelloway, 1998; Schumacker & Lomax, 1996). The covariance terms added to the model are listed in Table 1. After adding the covariance terms, the GFI, AGFI, S-RMR, and RMSEA were 0.95, 0.94, 0.035, and 0.043, respectively. Thus, the indexes that were obtained were deemed adequate to treat the dimensions as distinct factors in the path analytic model. Lambda-x estimates and standard errors as obtained from the confirmatory factor analyses and the item means across the Turkey data set are included in Table A.1 in Appendix A.

Table 1*Covariance Terms Added into the Confirmatory Factor Analysis Model*

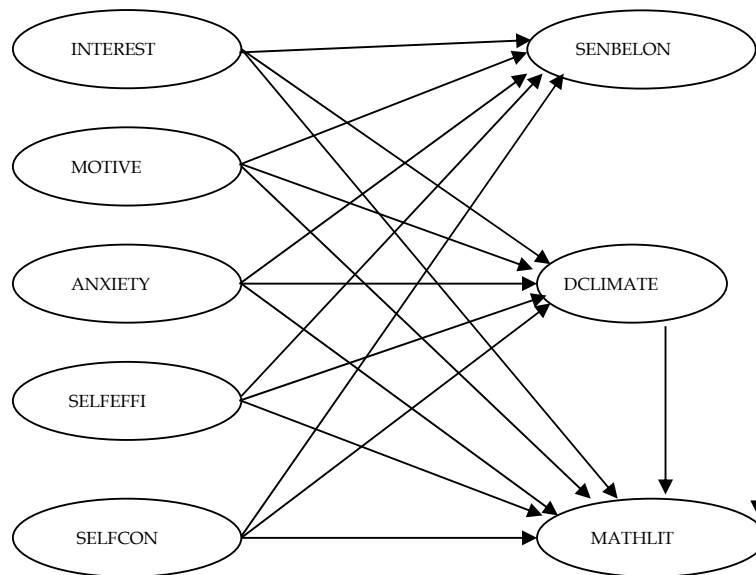
Covariance Terms Added Between Observed Variables		Values of Covariance Terms
Feel study-not good (notgood)	Feel study-worry (worry)	-0.18
Feel study-not good (notgood)	Feel study-tense (tense)	-0.14
Attitude further study (study)	Attitude career (career)	0.12
Feel study-not good (notgood)	Feel study-helpless (helpless)	-0.10
Feel study-nervous (nervous)	Feel study-worry (worry)	-0.12
Feel study-understanding difficult (difficul)	Feel study-quickly (quickly)	0.079
Feel study-poor marks (poormark)	Feel study-helpless (helpless)	0.12

Considering the factor structures obtained from the confirmatory factor analysis, seven latent variables that were parallel to the conceptual framework of the PISA Student Questionnaire were formed, i.e., *Interest in and Enjoyment of Mathematics*, *Instrumental Motivation in Mathematics*, *Anxiety in Mathematics*, *Self-Efficacy in Mathematics*, *Self-Concept in Mathematics*, *Sense of Belonging at School*, and *Disciplinary Climate in Mathematics*. In this particular analysis, the observed variable “Learning mathematics is worthwhile for me because it will improve my career prospects or chances” was loaded on *Interest in and Enjoyment of Mathematics* in addition to *Instrumental Motivation in Mathematics*. In Appendix A, Table A.1 indicates the names of the latent variables for each of the observed variables with their abbreviations and response modes.

The Cronbach alpha coefficients were 0.83, 0.82, 0.82, 0.82, 0.88, 0.68, and 0.79 for the latent variables *Interest in and Enjoyment of Mathematics*, *Instrumental Motivation in Mathematics*, *Anxiety in Mathematics*, *Self-Efficacy in Mathematics*, *Self-Concept in Mathematics*, *Sense of Belonging at School*, and *Disciplinary Climate for Mathematics Lessons*, respectively. The reliability coefficients of the dimensions were sufficiently high to treat them as distinct latent variables in the path analytic model.

The Hypothesized Model

Figure 1 illustrates the hypothesized model from the present study. *Interest in and Enjoyment of Mathematics (Interest)*, *Instrumental Motivation in Mathematics (Motive)*, *Anxiety in Mathematics (Anxiety)*, *Self-Efficacy in Mathematics (SelfEffi)*, and *Self-Concept in Mathematics (SelfCon)* were taken as exogenous variables, and *Sense of Belonging at School (SenBelon)*, *Disciplinary Climate for Mathematics Lessons (Dclimate)* and *Mathematical Literacy (Mathlit)* were taken as the endogenous variables. Among the endogenous variables, the paths from for *Sense of Belonging at School* and *Disciplinary Climate for Mathematics Lessons* to the *Mathematical Literacy* measure were also tested. In general, the model tests the impact of affective variables on both sense of belonging to school and climate in the classroom and consequently the direct and indirect effects of these variables on mathematical literacy measures.



Key: INTEREST - Interest in and Enjoyment of Mathematics; MOTIVE - Instrumental Motivation in Mathematics; ANXIETY - Anxiety in Mathematics; SELFEFFI - Self-Efficacy in Mathematics; SELFCON - Self-Concept in Mathematics; SENBELON - Sense of Belonging at School; DCLIMATE - Disciplinary Climate for Mathematics Lessons; MATHLIT - Mathematical Literacy

Figure 1: The Hypothesized Model

Results

When the significance of the standardized path coefficients was considered for the proposed model, observations indicated that the paths from *Self-Concept in Mathematics* to *Sense of Belonging at School*, from *Instrumental Motivation in Mathematics* to *Disciplinary Climate for Mathematics Lessons*, and from *Instrumental Motivation in Mathematics* to *Mathematical Literacy* were non-significant. Thus, these paths were deleted from the model. For the model that is illustrated in Figure 2, the GFI, AGFI, S-RMR, and RMSEA were 0.95, 0.94, 0.038, and 0.040, respectively. All of these fit indexes were deemed adequate to interpret the significant relationships among the latent variables. Lambda estimates, t-values and standard errors, Beta and Gamma estimates among the endogenous and exogenous variables, and the t-values are presented in Tables B.1 and B.2 in Appendix B.

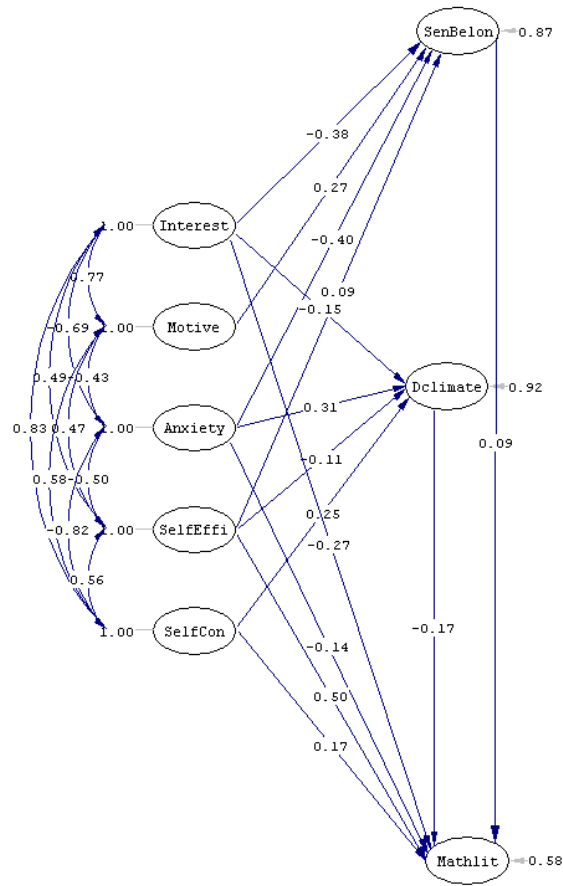


Figure 2: The Structural Model of Mathematical Literacy

As seen from Figure 2, the standardized path coefficients ranged between -0.40 and 0.50 in the fitted model. To interpret the magnitudes of the path coefficients, the criteria proposed by Cohen (1988) were used. Standardized path coefficients with absolute values less than 0.10 may indicate a “small” effect, whereas values around 0.30 indicate a “medium” effect; additionally, values above 0.50 indicate a “large” effect (Kline, 1998). With respect to these criteria, the path coefficient for *Self-Efficacy in Mathematics* to *Sense of Belonging at School*, the path coefficients for *Interest in and Enjoyment in Mathematics* and *Self-Efficacy in Mathematics* to *Disciplinary Climate for Mathematics Lessons*, and the path coefficients for *Anxiety in Mathematics*, *Self-Concept in Mathematics*, *Sense of Belonging at School* and *Disciplinary Climate in Mathematics Lessons* to *Mathematical Literacy* all indicate small effect sizes. In contrast, the path coefficient from *Self-Efficacy in Mathematics* to *Mathematical Literacy* could indicate a large effect size in the fitted model. All other path coefficients indicated medium effect sizes of varying magnitudes. The total variance in this particular model for the mathematical literacy measure is 42%.

In the fitted model, the greatest effect on *Mathematical Literacy* for the Turkish students came from *Self-Efficacy in Mathematics* with a total effect of 0.53. The other latent variables that also had strong effects on *Mathematical Literacy* were *Interest in and Enjoyment of Mathematics*, *Anxiety in Mathematics*, and *Disciplinary Climate for Mathematics Lessons* with total effects of -0.28, -0.23, and -0.17, respectively. On the other hand, *Self-Concept in Mathematics*, *Sense of Belonging at School*, and *Instrumental Motivation in Mathematics* indicated rather small relationships with the *Mathematical Literacy* measure with total effects of 0.13, 0.09, and 0.02, respectively. When the signs of the path coefficients were considered, observations indicated that *Self-Efficacy in Mathematics*, *Self-Concept in Mathematics*, *Sense of Belonging at School*, and *Instrumental Motivation in Mathematics* had positive relationships with the *Mathematical Literacy* measure, whereas *Interest in and Enjoyment of Mathematics*, *Anxiety in Mathematics* and *Disciplinary Climate for Mathematics Lessons* were negatively related to the *Mathematical Literacy* measure.

Discussion

In the present study, the analyses basically focused on the mathematical literacy of 15-year-old students as described using plausible values from the PISA 2003. In the structural model in question, *Sense of Belonging at School* and *Disciplinary Climate for Mathematics Lessons* were considered both exogenous and endogenous variables.

The latent independent variable *Self-Efficacy in Mathematics* was found to indicate the strongest total effect on the *Mathematical Literacy* measure. In the model under investigation, *Self-Efficacy in Mathematics* has a direct effect of 0.50 and a total effect of 0.53 on *Mathematical Literacy*. Self-efficacy in mathematics is defined as the degree to which students believe they are able to handle tasks effectively and overcome difficulties (OECD Publications, 2004a). This latent variable includes items that are concerned with the confidence required for students to successfully master specific learning tasks in line with mathematics related subjects (OECD Publications, 2004a, p. 136). The percentages of students who reported that they feel ‘very confident’ or ‘confident’ about the given proficiencies ranged from 25% to 42%. Thus, a number of

students seemingly do not feel confident in the proficiencies included in this particular latent variable. The strong relationship between this variable and mathematics performance is consistent with the results of previous research studies (Cooper & Robinson, 1991; Eccles et al., 1990; Ferry et al., 2000; Hackett & Betz, 1989; Hall & Ponton, 2005; O'Brien et al., 1999; OECD Publications, 2004a, 2005).

A surprising result was found in respect to the *Interest in and Enjoyment of Mathematics* latent variable; i.e., it was negatively related to the *Mathematical Literacy* measure. The relationship between interest in and enjoyment of mathematics and student performance in mathematics diminishes considerably or even becomes negligible in most countries when other learner characteristics are considered in the PISA project (OECD Publications, 2004a, p. 120, 121). However, no other instance reported a negative relationship in the OECD publications for this particular variable. This unexpected result might be due to various reasons. First, the PISA assesses a very unique construct as 'mathematical literacy'. This construct is not the learning achievement of the student per se. On the other hand, most of the studies reporting a positive relationship between attitudinal measures and the performance of the students consider learning achievement in mathematics (Abu-Hilal, 2000; Ferry et al., 2000; Ma, 1997; Reynolds & Walberg, 1992). Second, students who have positive attitudes towards mathematics might be successful at school in mathematics classes when learning achievement is concerned; however, no evidence to date supports a positive relationship between learning achievement and mathematical literacy as measured by the PISA project. A negative disposition may somehow foster students' efforts to succeed, and they might strive to be successful in the literacy items from the PISA 2003 assessment tools. Evaluations of the student endorsement ratios for the items in this particular latent variable revealed that 60% of students agree or strongly agree that they enjoy reading about mathematics, 58% agree that they do mathematics because they enjoy it, and 50% report that they look forward to their mathematics lessons (OECD Publications, 2004a).

Even though Turkish students have positive attitudes towards mathematics, they are not sufficiently proficient on the PISA 2003 literacy scale. As previously discussed, Turkish students performed very poorly on the PISA 2003. This low performance might be indicative of poor educational and social environments at schools that hinder learning in mathematics classes despite the positive interest and enjoyment reported by the students in reference to mathematics related activities. The negative relationship between *Interest in and Enjoyment of Mathematics* and the *Sense of Belonging at School* partly supports this argument since students with positive attitudes towards mathematics often do not feel a part of the school.

Even though the direct effect of *Interest in and Enjoyment of Mathematics* is negative on the *Mathematical Literacy* scale, it has a small positive indirect influence on the *Mathematical Literacy* measure over *Disciplinary Climate for Mathematics Lessons*. Students expressing more interest and enjoyment in learning mathematics are more likely to be orderly, quiet, listen to their teacher, and work well in their mathematics classes, thereby leading to a better performance for the mathematical literacy measure.

The latent variable of *Self-Concept in Mathematics* indicated a positive relation with the *Mathematical Literacy* measure. This particular variable has a direct effect of 0.17 and a total effect of 0.13 on literacy measure. The total effect decreases due to the negative

effect of *Disciplinary Climate for Mathematics Lessons* on the *Mathematical Literacy* measure. This finding is consistent with previous research studies as well (Abu-Hilal, 2000; Marsh, 1986). In the sample, 59% of the students agree or strongly agree that they are just not good at mathematics; 55% agree that they learn mathematics quickly; 53% agree that they get good marks in mathematics; 46% report that they have always believed that mathematics is one of their best subjects; and only 30% report that they understand even the most difficult work in their mathematics class (OECD Publications, 2004a). These percentages indicate that Turkish students need to develop more confidence in line with mathematics related issues. A comparison of the endorsement ratios of this particular dimension with the items in the *Interest in and Enjoyment of Mathematics* latent variable indicates that the Turkish students reported interest in and enjoyment of mathematics but still have confidence problems with mathematics related issues.

Anxiety in Mathematics has a negative direct effect of -0.14 and a total effect of -0.23 on *Mathematical Literacy* as expected. This result is supported by previous studies that have reported a significant and negative relationship between anxiety in mathematics and mathematics performance (Baloglu, 2004; Eccles et al., 1990; OECD Publications, 2004a, 2005). In the Turkish sample, 68% of the students agree or strongly agree that they worry about getting poor marks in mathematics; 64% agree that they often worry that mathematics classes will be difficult for them; 50% report that they get very tense when they have to do mathematics homework; 46% report that they feel helpless when doing a mathematics problem; and 41% agree that they get very nervous doing mathematics problems (OECD Publications, 2004a). Based on a consideration of the items from both the self concept and anxiety dimensions together, the Turkish students encounter great difficulty in achieving a good mark in mathematics; consequently, they worry about it. Grading seemingly becomes a basic source of anxiety among the students, which in general negatively affects their academic self confidence and increases their anxiety level in mathematics classes.

In the model tested, *Disciplinary Climate for Mathematics Lessons* has a direct effect on *Mathematical Literacy* with a -0.17 path coefficient. This result is supported by the findings of previous studies (Bos & Kuiper, 1999; Brookover et al., 1979; Scheerens & Bosker, 1997; Willms, 1992). As students become more orderly and quiet while listening to their teacher and working well in the classroom, they perform better on mathematical literacy skills. In the sample, 35 % of the students report that the teacher has to wait a long time for students to quiet down; 33% of the students report that noise and disorder are often present in the classroom; 31% of the students report that students cannot work well and that students do not start working for a long time after the lesson begins; and 24% of students report that students do not listen to what the teacher says (OECD Publications, 2004a). This data implies that a considerable classroom management problem is present in Turkish schools.

In the model under investigation, *Sense of Belonging at School* only shows a direct effect on *Mathematical Literacy*, and this effect is positive (i.e., 0.09 total effect). This particular latent variable has the lowest reliability estimate among the other latent variables used in the model. In the sample, 41% of the students agree or strongly agree that other students seem to like them; 25% agree that they feel lonely; 14% agree that they feel like an outsider or left out of things; and 11% report that they feel awkward and out of place (OECD Publications, 2004a).

Instrumental Motivation in Mathematics has the smallest total effect on *Mathematical Literacy* in the model under investigation. Instrumental motivation is beyond a general interest in mathematics and refers to how 15-year-olds assess the relevance of mathematics to their own life and what role such external motivation plays with regard to their mathematics performance (OECD Publications, 2004a). In Turkey, 86% of the students agree or strongly agree that learning mathematics is important because it will help them with the subjects that they want to study further in school; 81% report that making an effort in mathematics is worthwhile because it will help them in the work that they want to do later; 79% agree or strongly agree that mathematics is an important subject for them because they need it for what they want to study later; and 66% report that they will learn many things in mathematics that will help them get a job (OECD Publications, 2004a). This data indicates a favorable response pattern by Turkish students for this particular latent variable. However, the indirect effect of this variable on *Mathematical Literacy* through *Sense of Belonging to School* is too small to interpret (i.e., 0.02).

When the relationships between the endogenous variables and other exogenous variables in the model were considered, the strongest effect on *Sense of Belonging at School* came from *Anxiety in Mathematics*. The students who are less anxious about mathematics feel a greater sense of belonging at school. This finding was expected due to the positive relationship between sense of belonging at school and the mathematical literacy performances of the students (OECD Publications, 2004a). The other two latent independent variables that have positive relationships with *Sense of Belonging at School* are *Instrumental Motivation in Mathematics* and *Self-Efficacy in Mathematics* with the total effects of 0.27 and 0.09, respectively. This is also an expected outcome; specifically, as the positive disposition of students increases in terms of motivation related factors and self efficacy, they increasingly feel a part of the school to which they belong (Turanlı, 2009). *Interest in and Enjoyment of Mathematics* indicates a direct effect of -0.38 with the *Sense of Belonging at School* latent variable. This indicates that students with positive attitudes towards mathematics do not feel a part of the school. This finding could be explained by the social and academic environment that schools provide for the students as previously discussed.

The factor that has the strongest total effect on *Disciplinary Climate for Mathematics Lessons* is *Anxiety in Mathematics* with an effect of 0.31. Students who are less anxious about mathematics typically become more orderly and quiet and listen to their teacher while working well in the classroom. *Self-Concept in Mathematics* is the second strongest factor affecting *Disciplinary Climate for Mathematics Lessons* with a total effect of 0.25. This result indicates that a higher level of self-concept somehow creates a classroom management problem in Turkish schools. Seemingly, a higher level of belief about one's academic ability in mathematical related tasks makes students behave in a less

disciplined manner in the classroom. The positive relation between self concept and mathematics literacy implies that these students have a higher level of academic achievement.

Interest in and Enjoyment of Mathematics and *Self-Efficacy in Mathematics* have negative direct effects of -0.15 and -0.11, respectively, on *Disciplinary Climate for Mathematics Lessons*. These results were expected. Therefore, higher confidence in mathematics related tasks and a positive attitude towards mathematics provide a more disciplined environment in the classroom.

In the present study, the 42% total variance for the mathematical literacy measure indicates the importance of affective variables in explaining the academic performance of the students. Turkish students have positive attitudes towards mathematics with less confidence and higher anxiety levels. Somehow, the positive attitude of the students did not result in a better academic performance in the educational system. Moreover, receiving good marks in mathematics classes is a major source of anxiety and low self concept. Additionally, evidence indicates that the classroom climate is negatively influenced by this high anxiety and lack of confidence on the part of the students. This might cause classroom management problems for teachers dealing with students who perform on different academic levels and have various backgrounds in mathematics.

Educational policy makers in Turkey may need to consider the student related affective variables, such as academic self concept, anxiety, etc., in order to enhance the quality of their educational practices via pre-service and in-service teacher training programs and school curricula and by emphasizing the classroom management skills of a successful teacher. On the other hand, the negative direct effect of interest dimension on mathematical literacy indicates a need for the closer investigation of the attitudinal characteristics of the students and their impact on cognitive measures in further research studies and analyses of the international data sets.

Appendix A

Table A. 1 Latent and Observed Variables, Abbreviations for Latent Variables, Response Modes & LISREL Estimates, Standard Errors for Confirmatory Factor Analysis and Item Means

Observed Variables	Latent Variables	Response Modes	Lambda-x	SE	Mean
I do Mathematics because I enjoy it			0.94	0.11	2.68
I look forward to my Mathematics lessons	Interest in &	1; strongly	0.91	0.17	2.52
I enjoy reading about Mathematics	Enjoyment	disagree	0.80	0.35	2.70
Learning Mathematics is worthwhile for me because it will improve my career prospects, chances	of Mathematics (INTEREST)	4; strongly agree	-0.23		
Learning Mathematics is worthwhile for me because it will improve my career prospects, chances			0.95	0.39	3.20
Making an effort in Mathematics is worth it because it will help me in the work that I want to do later on	Instrumental Motivation in	1; strongly disagree	0.87	0.25	3.09
Mathematics is an important subject for me because I need it for what I want to study later on	Mathematics (MOTIVE)	4; strongly agree	0.80	0.36	3.07
I get very tense when I have to do Mathematics homework			0.82	0.28	2.49
I often worry that it will be difficult for me in Mathematics classes	Anxiety in	1; strongly	0.82	0.28	2.76
I get very nervous doing Mathematics problems	Mathematics	disagree	0.74	0.45	2.36
I feel helpless when doing a Mathematics problem	(ANXIETY)	4; strongly agree	0.65	0.54	2.45
I worry that I will get poor marks in Mathematics			0.54	0.71	2.84

Appendix A (Continued)

Table A. 2 Latent and Observed Variables, Abbreviations for Latent Variables, Response Modes & LISREL Estimates, Standard Errors for Confirmatory Factor Analysis and Item Means (Continued)

Observed Variables	Latent Variables	Response Modes	Lambda-x	SE	Mean
Using a train timetable to work out how long it would take to get from one place to another			0.78	0.38	2.72
Calculating how much cheaper a TV would be after a 30% discount			0.77	0.41	2.97
Calculating how many square meters of tiles you need to cover a floor	Self-Efficacy in	1; not at all	0.76	0.42	2.84
Calculating the petrol consumption rate of a car	Mathematics	confident	0.68	0.54	2.61
Understanding graphs presented in newspapers	(SELFEFFI)	4; very confident	0.62	0.61	2.86
Finding the actual distance between two places on a map with a 1:10,000 scale			0.62	0.62	2.77
I have always believed that Mathematics is one of my best subjects			0.90	0.19	2.46
I get good marks in Mathematics	Self-Concept	1; strongly	0.83	0.31	2.56
I learn Mathematics quickly	In Mathematics	disagree	0.81	0.34	2.61
I am just not good at Mathematics ^a	(SELFCON)	4; strongly agree	0.81	0.33	2.36
In my Mathematics class, I understand even the most difficult work			0.74	0.46	2.17
I feel like an outsider or left out of things ^a	Sense of Belonging		0.75	0.43	3.28
I feel lonely ^a	at School	1; strongly	0.73	0.46	3.01
I feel awkward and out of place ^a	(SENBELON)	disagree	0.69	0.53	3.38
Students don't start working for a long time after the lesson begins			0.79	0.38	2.17
There is noise and disorder	Disciplinary Climate	1; never or	0.79	0.37	2.26
Students cannot work well	for Mathematics	hardly	0.69	0.53	2.24
The teacher has to wait a long time for students to quieten down	Lessons	ever	0.68	0.54	2.27
Students don't listen to what the teacher says	(DCLIMATE)	4; every lesson	0.62	0.62	2.14

^aThese items were reversed in the analyses.

Appendix B

Table B. 1 LISREL Estimates, T-Values, and Standard Errors for LISREL Model

Latent Variables	Observed Variables	Lambda	T-Values	SE
INTEREST	Attitude enjoy Maths	0.94	73.87	0.01
	Attitude look forward	0.91	69.84	0.01
	Attitude enjoy reading	0.80	57.07	0.01
	Attitude career	-0.23	-9.20	0.02
MOTIVE	Attitude career	0.94	32.41	0.03
	Attitude effort	0.87	58.76	0.01
	Attitude further study	0.80	52.73	0.02
ANXIETY	Feel study-tense	0.83	60.84	0.01
	Feel study-worry	0.82	59.05	0.01
	Feel study-nervous	0.74	48.66	0.02
	Feel study-helpless	0.65	42.85	0.02
	Feel study-poor marks	0.54	33.52	0.02
SELFEFFI	Confident timetable	0.79	53.29	0.01
	Confident discount	0.78	52.34	0.01
	Confident area	0.76	50.98	0.01
	Confident rate	0.67	42.73	0.02
	Confident graphs	0.63	39.25	0.02
	Confident distance	0.62	38.43	0.02
SELFCON	Feel study-best subject	0.90	67.64	0.01
	Feel study-good marks	0.83	59.89	0.01
	Feel study-quickly	0.81	57.37	0.01
	Feel study-not goodb	0.81	58.78	0.01
	Feel study-underst. diffc.	0.74	49.77	0.01
SENBELON	Feel an outsider ^b	0.76	43.04	0.02
	Feel lonely ^b	0.73	41.58	0.02
	Feel awkward ^b	0.68	39.98	0.02
DCLIMATE	Lesson late start	0.79	51.49	0.02
	Lesson noise	0.79	51.44	0.02
	Lesson can't work well	0.69	42.91	0.02
	Lesson quieten down	0.68	42.78	0.02
	Lesson don't listen	0.62	37.70	0.02
MATHLIT	Plausible value 1	0.95	69.69	0.09
	Plausible value 2	0.95	69.57	0.09
	Plausible value 3	0.95	69.34	0.10
	Plausible value 4	0.95	69.73	0.09
	Plausible value 5	0.95	69.59	0.09

^bThese items were reversed in the analyses.

Appendix B (Continued)**Table B. 2 Beta and Gamma Estimates and T-Values for LISREL Model**

Latent Variables (from)	Latent Variables (to)	Beta	Gamma	T-Value
Sense of Belonging at School	Mathematical Literacy	0.09	-	5.55
Disciplinary Climate for Mathematics Lessons	Mathematical Literacy	-0.17	-	-10.68
Interest in and Enjoyment of Mathematics	Sense of Belonging at School	-	-0.38	-8.23
Instrumental Motivation in Mathematics	Sense of Belonging at School	-	0.27	6.87
Anxiety in Mathematics	Sense of Belonging at School	-	-0.40	-12.38
Self-Efficacy in Mathematics	Sense of Belonging at School	-	0.09	3.41
Interest in and Enjoyment of Mathematics	Disciplinary Climate for Mathematics Lessons	-	-0.15	-3.88
Anxiety in Mathematics	Disciplinary Climate for Mathematics Lessons	-	0.31	8.80
Self-Efficacy in Mathematics	Disciplinary Climate for Mathematics Lessons	-	-0.11	-4.61
Self-Concept in Mathematics	Disciplinary Climate for Mathematics Lessons	-	0.25	4.94
Interest in and Enjoyment of Mathematics	Mathematical Literacy	-	-0.27	-9.11
Anxiety in Mathematics	Mathematical Literacy	-	-0.14	-5.06
Self-Efficacy in Mathematics	Mathematical Literacy	-	0.50	23.45
Self-Concept in Mathematics	Mathematical Literacy	-	0.17	4.36

References

- Abu-Hilal, M. M. (2000). A structural model for predicting mathematics achievement: Its relation with anxiety and self-concept in mathematics. *Psychological Reports, 86*, 835-847.
- Baloğlu, M. (2004). The relationship between different ways of coping and mathematics anxiety. *Eğitim Araştırmaları - Eurasian Journal of Educational Research, 16*, 95-101.
- Bos, K., & Kuiper, W. (1999). Modeling TIMSS data in a European comparative perspective: Exploring influencing factors on achievement in mathematics in grade 8. *Educational Research and Evaluation, 5*(2), 157-179.
- Brookover, W. B., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *School social systems and student achievement: Schools can make a difference*. New York: Praeger.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models*. (p.136-162). Newbury Park, CA: Sage.
- Cooper, S. E., & Robinson, D. A. G. (1991). The relationship of mathematics self-efficacy beliefs to mathematics anxiety and performance. *Measurement & Evaluation in Counseling & Development, 24*(1).
- Dorman, J. P. (2001). Associations between classroom environment and academic efficacy. *Learning Environments Research, 4*, 243-257.
- Eccles, J. S., Meece, J. L., & Wigfield, A. (1990). Predictors of mathematics anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*(1), 60-70.
- Ferry, T. R., Fouad, N. A., & Smith, P. L. (2000). The role of family context in a social cognitive model for career-related choice behavior: A mathematics and science perspective. *Journal of Vocational Behavior, 57*, 348-364.
- Fraser, B. J. (1986). *Classroom environment*. London: Croom Helm.
- Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of the goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models*. (p.40-65). Newbury Park, CA: Sage.
- Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy / mathematics performance correspondence. *Journal for Research in Mathematics Education, 20*, 263-271.
- Hall, J. M., & Ponton, M. K. (2005). Mathematics self-efficacy of college freshman. *Journal of Developmental Education, 28*(3), 26-33.
- Hill, P. W., & Rowe, K. J. (1998). Modeling educational effectiveness in classrooms: The use of multi-level structural equations to model students' progress. *Educational Research and Evaluation, 4*(4), 307-347.
- İş Güzel, Ç., & Berberoğlu, G. (2005). An analysis of the Programme for International Student Assessment 2000 (PISA 2000) mathematical literacy for Brazilian, Japanese and Norwegian students. *Studies in Educational Evaluation, 31*, 283-314.
- Jöreskog, K., & Sörbom, D. (1993). *Structural equation modeling with the SIMPLIS command language*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Jöreskog, K., & Sörbom, D. (1999). *LISREL 8.30*. Chicago: Scientific Software International Inc.

- Kelloway, E. K. (1998). *Using LISREL for structural equation modeling*. London, New Delhi: Sage Publications.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: The Guilford Press.
- Ma, X. (1997). Reciprocal relationships between attitude toward mathematics and achievement in mathematics. *Journal of Educational Research*, 90, 221-229.
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23(1), 129-149.
- National Education Publications (2005). *National report of PISA 2003*. Ankara: National Education Publications.
- O'Brien, V., Martinez-Pons, M., & Kopala, M. (1999). Mathematics self-efficacy, ethnic identity, gender, and career interests related to mathematics and science. *The Journal of Educational Research*, 92(4), 231-235.
- Organisation for Economic Co-Operation and Development (2003). *The PISA 2003 assessment framework - Mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD Publications.
- Organisation for Economic Co-Operation and Development (2004a). *Learning for tomorrow's world*. Paris: OECD Publications.
- Organisation for Economic Co-Operation and Development (2004b). *PISA 2003 Web Site*. Available online at: <http://www.pisa.oecd.org> (accessed first on 05 January 2005).
- Organisation for Economic Co-Operation and Development (2005). *PISA 2003 technical report*. Paris: OECD Publications.
- Papanastasiou, C. (2000). Effects of attitudes and beliefs on mathematics achievement. *Studies in Educational Evaluation*, 26, 27-42.
- Papanastasiou, C. (2002). Effects of background and school factors on the mathematics achievement. *Educational Research and Evaluation*, 8(1), 55-70.
- Papanastasiou, E. (2002). Factors that differentiate mathematics students in Cyprus, Hong-Kong and the USA. *Educational Research and Evaluation*, 8(1), 129-146.
- Reynolds, A. J., & Walberg, H. J. (1992). A structural model of high school mathematics outcomes. *Journal of Educational Research*, 85(3), 150-158.
- Scheerens, J., & Bosker, R. J. (Eds.) (1997). *Foundations of educational effectiveness*. London: Routledge.
- Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Schreiber, J. B. (2002). Institutional and student factors and their influence on advanced mathematics achievement. *Journal of Educational Research*, 95(5), 274-286.
- Turanlı, A. S. (2009). Students' perceptions of teachers' behaviors of socioemotional support and students' satisfaction with the classroom atmosphere. *Egitim Arastirmalari - Eurasian Journal of Educational Research*, 35, 129-146.
- Willms, J. D. (1992). *Monitoring school performance: A guide for educators*. Washington, DC: Falmer.

**Uluslararası Öğrenci Değerlendirme Programı'nda (Pisa 2003)
Öğrencilerin Duyuşsal Özellikleri ve Bu Özelliklerin Matematik Okur
Yazarlığı İle İlişkisi
(Özet)**

Problem Durumu: Bu çalışmada, Uluslararası Öğrenci Değerlendirme Programı'nda (PISA 2003) öğrencilerin matematik okur yazarlıkları ile ilişkili duyuşsal faktörler değerlendirilmektedir.

Araştırmanın Amacı: Bu çalışmanın amacı, yapısal eşitlik modeli kullanılarak 15 yaşındaki Türk öğrencilerin Uluslararası Öğrenci Değerlendirme Programı'ndaki (PISA 2003) matematik okur yazarlıkları ile ilişkili duyuşsal değişkenlerin incelenmesidir.

Araştırmanın Yöntemi: PISA 2003 matematik okur yazarlığı ile ilişkili olması beklenen duyuşsal değişkenler ele alınarak PISA Türkiye verileri yapısal eşitlik modellemesi çerçevesinde analiz edilmiştir. Bu anlamda, matematik okur yazarlığı ile ilişkili olarak incelenen örtük değişkenler; matematiğe yönelik ilgi, matematiğe yönelik motivasyon, matematiğe ilişkin kaygı, matematikte kendini yeterli görme, matematikte özgüven, okula ait olma duygusu ve matematik derslerindeki sınıf disiplini ve ortamı olarak sıralanmaktadır. Yapısal eşitlik modelinde okula ait olma duygusunun ve matematik derslerindeki sınıf disiplini ve ortamının da matematik okur yazarlığı ile ilişkisi de değerlendirilmektedir. Bu anlamda, okula ait olma duygusu ve matematik derslerindeki sınıf disiplini ve ortamı hem bağımsız hem de bağımlı örtük değişkenler olarak ele alınmaktadır.

Araştırmanın Bulguları: Matematik okur yazarlığı ile en büyük ilişkiyi veren örtük değişken matematikte kendini yeterli görme olarak karşımıza çıkmaktadır. Matematik okur yazarlığı ile istatistiksel olarak anlamlı ilişki veren diğer örtük değişkenler, matematiğe yönelik ilgi, matematiğe ilişkin kaygı ve matematik derslerindeki sınıf disiplini ve ortamı olarak gözlenmektedir. Matematiğe yönelik ilgilerini yüksek olarak belirten öğrencilerin matematik okur yazarlık performanslarının daha düşük olduğu bulunmaktadır. Aksine, matematiğe yönelik ilgi değişkeninin, matematik okur yazarlığına, matematik derslerindeki sınıf ortamı değişkeni üzerinden küçük ancak pozitif olan bir dolaylı ilişkisi de gözlenmektedir.

Araştırmanın Sonuçları: Bu çalışmada, yapısal eşitlik modeli ile matematik okur yazarlık varyansının %42'si açıklanmaktadır. Açıklanan varyans miktarı da, öğrencilerin akademik performanslarının açıklanmasında duyuşsal özelliklerin ne denli önemli olduğunu ortaya koymaktadır. Türk öğrencilerin genelde matematiğe yönelik olumlu tutumlara sahip olduğu, ancak matematikte özgüven düzeylerinin düşük, matematiğe ilişkin kaygı düzeylerinin de yüksek olduğu görülmektedir. Öğrencilerin bu olumlu tutumları eğitim sistemi içerisinde bir şekilde daha yüksek akademik başarı getirecek biçimde yönlendirilememektedir. Araştırmada, ayrıca, kaygı ile özgüven düzeylerinin artmasının sınıf ortamını negatif olarak etkilediği ile ilgili bulgulara da ulaşılmaktadır. Bu anlamda, sınıflarda bazı sınıf yönetimi sorunlarının yaşandığına dair bir bulgu olarak da yorumlanabilir.

Anahtar Kelimeler: Öğrenci Başarı Değerlendirme Programı (PISA); Matematik okur yazarlığı; Doğrulayıcı faktör analizi; Yapısal eşitlik modeli; Türkiye

Purposes and Identities of Higher Education Institutions, and Relatedly the Role of the Faculty

Zeynep Kızıltepe*

Suggested Citation:

Kızıltepe, Z. (2010). Purposes and identities of higher education institutions, and relatedly the role of the faculty. *Egitim Arastirmalari - Eurasian Journal of Educational Research*, 40, 113-132.

Abstract

Problem Statement: The purposes of Higher Education Institutions (HEIs), their identities, the knowledge transmitted, and the roles of the academicians working in those institutions have been investigated recently by several researchers. HEIs are considered as if they have outlived their usefulness in the classic sense, because knowledge has acquired a more utility-oriented meaning today than culture-oriented. Modern beliefs about their purposes caused universities to steer towards functional curriculums, downplaying the classic ones. With its policies, programs, and practices designed to help students prepare for a better job or meet state, regional, and local economic needs, higher education is now perceived as a private benefit rather than as a public good.

Purpose of the Study: This paper investigates the outcomes of a study describing the perspectives of academicians working in the faculty of education of a state university about the purposes and identities of HEIs, the knowledge that is created and transmitted in those institutions, and the roles of the faculty members in relation to the above-mentioned issues.

Methods: One-to-one semi-structured interviews were conducted with eighteen university academicians, all members of the faculty of education of a state university in Istanbul, after they were pre-tested on a sample of three people. The faculty members were interviewed by the author only, and the duration of each interview varied from forty to eighty-five minutes. Interviews were recorded and transcribed verbatim, then subjected to close examination. The analysis of the data employed was

* Associate Prof. Dr., Boğaziçi University Faculty of Education,
zeynep.kiziltepe@boun.edu.tr

associated with the constant comparative method, to identify recurring themes within and across data sources.

Results: Although a variety of viewpoints were found, the respondents agreed that major purposes of higher education institutions should be to promote “active citizenry” among their students. In addition to preparing students for the workplace, the knowledge that is created and shared should also contribute to the well-being and prosperity of humanity. Of special significance was identification of the widely held view that all higher education institutions should have an identity to be internalized by the academicians – the most important figures in the system – who pass it on to their students to become responsible citizens in that society.

Conclusions and Recommendations: Without determining the purposes of HEIs, plans about what to learn are meaningless. In order to remain viable actors in a changing political, social, and economic environment, universities must redefine their roles in their societies. The purpose of higher education is not only the development of people, or of society, or of economy but people, society, and the economy all together; furthermore, higher education, in turn, develops with these three themes in mind. This requires interpersonal competence, multi-cultural understanding, skills in problem identification and problem solving, a sense of purpose, and the confidence that you can act in ways that make a difference.

Keywords: Motivation, demotivation, faculty, purposes of the university, identity of the university.

Over the last several decades, a body of research has been developed investigating the purposes of Higher Education Institutions (HEIs), their identities, the knowledge transmitted, and the role of the faculty working in those institutions (Akerlind, 2004; Bleiklie & Byrkjeflot, 2002; Lutz & Field, 1998). HEIs are considered as if they have outlived their usefulness in the classic sense, because knowledge has acquired a more encompassing meaning today, being more utility-oriented than culture-oriented (Bleiklie, 2005; Gibbons, 1998). However, several perspectives are in favor of the constructivist duty of the universities rather than their utilitarian merits (Erzinçlioğlu, 1999; Nixon, 1996). While some suggest that universities can play a key role in modernity by expanding the discursive capacity of society and enhancing citizenship in the knowledge society (Delanty, 2001), others believe that “simplistic measures of monetary success, social mobility, or grade point averages are insufficient measures of determining the total value higher education should play in the overall human development of individuals” (Kempner & Taylor, 1998, p. 302). However, modern beliefs about their purposes caused universities to steer towards functional curriculums, downplaying the classic ones. With its policies, programs, and practices designed to help students prepare for a better job or meet state, regional, and local economic needs, higher education is now perceived as a private benefit rather than as a public good (Chickering, 2002, p. 5). Similarly, Chomsky

(2000) claims that universities are economically parasitic now, relying on external support instead of being self-sustaining.

This change, supported by national governments and university leaders, is based on the assumption that HEIs are “market or quasi-market organizations striving to become entrepreneurial in their approach to teaching and research” (Bleiklie, 2005, p. 31). Furthermore, he points out that HEIs have a more global reach now, with the attempt to turn the institutions into dynamic, entrepreneurial, high-quality enterprises. They are not only affected by national or international organizations but supranational ones as well, such as the Organization for Economic Co-operation and Development (OECD), United Nations Educational, Scientific and Cultural Organization (Unesco), and World Trade Organization (WTO).

The relationship between the faculty members and students has also changed, reflecting a model in which the student is cast as customer or client and the teacher as service deliverer (Rowland, 2002). Some even believe that by treating students as customers, education as a service product, and applying the lessons of service marketing to the classroom, student satisfaction is guaranteed (McCollough & Gremier, 1999). Student satisfaction is believed to be one potential way to improve instructor performance, increase accountability in the classroom, and deal with the characteristics of high intangibility, high coproduction, and heterogeneity characteristics of education (Özgüngör, 2010). However, a university degree today does not give its holder the proof that s/he has acquired a full capacity of learning or understanding (Erzinçlioğlu, 1999).

The reasons for this drastic change from higher education to professional education are various. At the beginning, the liberal and the practical stood side by side embracing the traditional notions of higher education such as intellectual, moral, and civic values, and the newer expectations which encompass the idea that learning is practical and vocational and that the state institutions had a responsibility to serve public needs (Eddy, 1957). However, the knowledge revolution, which took place around the turn of the millennium, has shifted the nature of work towards occupations associated with knowledge and information, rather than those rooted in industrial production (Grubb & Lazerson, 2005). Having greater emphasis on vocational purposes, higher education has come to emphasize its private benefits for individuals, such as access to income and professional status undermining moral, civic, and intellectual purposes of education. Along the same line, Bryce and Humes (2000, p. 1011) state that “Knowledge is now seen as a commodity... with a market value... and learners are human capital which it is in the countries’ interest to invest”. Gibbons (1998), on the other hand, believes that massification of higher education and the pressures of international competition have together contributed to the change.

Similarly, due to developments in globalization, and information technologies, the education sector in the developed world has changed, both in shape and content (Gültekin, Çelik, & Nas, 2008). In the face of the influences of marketisation and globalisation, students are becoming increasingly apolitical, apathetic, instrumental,

consumerist, competitive, calculating, pragmatic and job-oriented, or “lacking in passion”, as several Russians put it (Rowland, 2002, p. 55). In the U.S.A., higher education is claimed to be the clearest embodiment of the American dream of getting ahead, especially getting ahead through one's own labor (Lazerson, 1998). The bottom line is students are no longer individuals seeking upward mobility; they are future workers (Rhoades & Slaughter, 1997).

Given these analyses, the present study had four purposes: (a) to explore how Turkish academicians perceive the purposes of HEIs, (b) to inquire if HEIs should have identities and how these identities are and/or should be established, (c) to find what kind of knowledge is offered in those institutions, and (d) to understand what this group believes the role of the faculty is in connection with the above issues. Therefore the research questions are

1. What are the purposes of HEIs in a society?
2. What kind of knowledge is offered in HEIs?
3. Do HEIs have identities?
4. What are the roles of the faculty members in HEIs?

Method

The present study was carried out in a public university in Istanbul, Turkey, which had 11,694 student enrollments in both undergraduate and graduate level programs in the academic year 2009-2010 (Erkoç, 2010). The university has four faculties, namely The Faculty of Arts and Sciences, The Faculty of Economics and Administrative Sciences, The Faculty of Engineering, and The Faculty of Education. This study was conducted in the Faculty of Education, which has five main departments, namely Computer Education and Educational Technology, Educational Sciences, Foreign Language Education, Primary Education, and Secondary School Science & Mathematics Education.

Sample

The study was done in the spring of 2008, with eighteen Turkish faculty members, who were chosen by convenience sampling. The specific information about the faculty members is listed below:

- *Discipline*: Six from educational sciences, six primary school education, three foreign language education, three secondary school science and mathematics education.
- *Academic Experience*: Two years to sixteen years.
- *Appointment*: Three tenured associate professors, ten assistant professors, five instructors with PhDs.
- *Gender*: All women.
- *Age range*: Mid-20s to early 50s.

The Design and Procedure

The outcomes presented in this exploratory study are based on one-to-one semi-structured interviews. The reason for choosing a qualitative design is that “it implies a direct concern with experience as it is lived or felt or undergone” (Sherman & Webb, 1988, p. 7). Furthermore, qualitative studies are interested in understanding how people make sense of the world and the experiences they have in the world (Merriam, 1998). Since the aim of this study is to investigate the perceptions and beliefs of the individual faculty members of the purposes, identities of HEIs, and the role of the faculty in these institutions, the study employed semi-structured face-to-face interviews. Semi-structured face-to-face interviews are planned, prearranged interactions between two or more people, and their purpose is to generate detailed and desired information about an event, that would not otherwise be possible to obtain by other means of data collection (Lankshear & Knobel, 2006, p. 198).

The open-ended interview questions were pre-tested on a sample of three people. They were prepared in Turkish and were conducted in Turkish. However, the answers were translated into English for this manuscript and the translations were edited by a native speaker of English. Visits to the interviewees’ offices were made by appointment. They were interviewed by the author only, who is an assistant professor in the same faculty, and were assured that their responses would not be for attribution. They were asked and granted permission for taping of the interviews. The duration of each interview varied from forty to eighty-five minutes. The interviews were transcribed and coded using Seventh String Software version 7.50.0 for Mac. The interview questions were:

1. What do you think the purposes of HEIs are in a society?
2. What kind of knowledge do you think is offered in these institutions?
3. What are the roles of the faculty members in these institutions?
4. Do you think HEIs have identities? If either “yes” or “no”. Why do you think so?

During the interviews, unstructured follow-up questions were used to elicit more information on a given topic, or to clarify some words or phrases that the interviewees had used. These questions commonly took the form of “Could you explain that a bit more?” “What do you mean by that?” “Could you give me an example?” and, “Why do you think this is the situation?” However, the interviewer was careful not to give her own ideas on the topic.

Interviews were recorded and transcribed verbatim, then subjected to close examination. The analysis of the data employed was associated with the constant comparative method, to identify recurring themes within and across data sources (Merriam, 1998). It is a widely used method in all kinds of qualitative studies (Glaser & Strauss, 1967). The basic strategy is beginning with a particular incident from an interview and comparing it with another incident in the same set of data or in another set. These comparisons lead to tentative categories, which are then compared to each other until thematic units are reached (Merriam, 1998). This meant reading

the texts several times, comparing and contrasting them for similarities and differences, and aggregating them into themes.

Results and Discussion

In this study, four themes emerged as the results of the analysis: (i) the purposes of the universities; (ii) the identities of the universities; (iii) the knowledge transmitted in these institutions; and (iv) the roles of the faculty members. Interestingly, they were found to be all interrelated and closely interconnected. Furthermore, there was a consensus of opinion on them among the academicians interviewed. The categories are summarized in detail below, with verbatim quotes from the original transcripts used to support the key aspects.

The Purposes Of Higher Education Institutions

In this category, it was generally stated that the purpose of HEIs is not only to inform students about a certain topic, but at the same time to shape their lives, their thinking and life styles. Universities are places for civic learning and socially responsible behavior. It is the place where students start enlarging their perspectives.

Universities are not places students are content with what is offered to them as knowledge but start adding to what is offered. They are places where one gets not only the technical knowledge about one's own field of study and graduates but they are places where one starts thinking widely and in-depth about a lot of things such as life, society, beliefs.
(Primary Education)

However, most interviewees believed that universities recently have been converted to higher-level vocational schools with their departmentalized programs. They claimed that this is encouraged by the capitalist system dominant in societies around the world. This, in turn, has considerable effects on the purposes of HEIs.

The purpose of HEI in Turkey is unfortunately, I deliberately say unfortunately and underline it, to prepare students for the corporate world. I say unfortunately because in my opinion the main aim of HEI is to prepare the individual as a responsible citizen, to raise the future leaders. When I say leader, it is not just a leader in the business world, [but] a leader in the social life or in the family too. But the opposite is done in Turkey. Students are educated in specialized, departmentalized areas without taking any liberal arts courses. They graduate from faculties of economics or administrative sciences or law or education without taking these courses, and they miss the philosophies behind their careers. In my opinion, HEI are more and more preparing students to meet the needs of the corporate world.

Why do you think that is the situation?

Because money is important now. We are living in a capitalist system. Economic values are overrated now; ethical values are forgotten. Why

is that so? The primary concern now is to find a job so that the person can bring bread to the family. This is more crucial in developing countries like Turkey. And how is a job found better? If you have the qualifications. You have to have a bachelor's degree, or better a master's degree from a university so that you can get ahead of others. Universities are now aware of these needs, and they are opening programs accordingly. (Educational Sciences)

It is worth noting here that the acknowledged purpose of higher education in the United Kingdom (U.K.) is to create wealth for a global economy (Rowland, 2002). Blackstone (2001) reports the U.K. minister of State for Higher Education describing higher education solely in terms of the needs of the global economy. Another research study (Rowland, 2001) confirmed that the situation is no different in South Africa and Russia, and a study from Costa Rica claims that the changing political, social, and economic climate, and the growing private university sector call into question the traditional social role of public higher education in that country (Twombly, 1997). Similarly, Kwiek (2003) reports that the higher education system in Poland has been undergoing dynamic changes since 1989. It has gone from an elite higher education system to mass higher education with a strong private sector connection, leaving public higher education with less support and funding by the government. However, the "liberal" element of a Bachelor's degree program aims to expose students to broader forms of knowledge, reasoning, and critical inquiry, and ideally to developing the skills needed for civic participation (Cornwell & Stoddard, 2001, p. 6).

Studies from Turkey reflect similar ideas. Universities which function nowadays just like vocational schools, give the impression that when one graduates from universities, studying, doing research, and learning are over (Çakmakçı, 2009). However, universities should provide the means for people to gain a culture, to contribute to the society by doing research, to catch the demands of the contemporary world or to be open to change (Demircioğlu, 2007; Tufan, 2008). It is claimed that the current expectations of HEIs should be listed as: to prepare qualified people in various fields of work, to enable people to gain a vocation, to produce knowledge by doing research, and to enlighten the society (Oğuz, 2004; Sönmez, 2003). Therefore, HEIs have always become a symbol of prestige in their countries (Kaya, 1989, p. 80).

The Knowledge Created, Offered, Or Used In Those Institutions

In this category, knowledge that is currently offered in HEIs is believed to be mostly technical; it is not a creative and experimental kind of knowledge that enables students to get the necessary cultural orientations. There is an increasing emphasis on vocational knowledge in the courses offered. The knowledge developed is in line with a market utility. Knowledge is no longer created for the sake of knowledge itself.

Knowledge is everywhere now but how shall I reach that knowledge? Because there is junk knowledge everywhere too. How shall I analyze this knowledge or discuss it thoroughly? This skill should be taught too, the skill to do research, the skill to think creatively. How can you be more intellectual? How can you have a vision? These are the most important things in a university. They are more important than the technical knowledge. (Secondary School Science and Mathematics Education)

Moreover, interviewees believed that what is taught in most universities in Turkey (and in some parts of the world) is the widely accepted knowledge that is not negotiable, not questionable. Teachers usually go into classes, lecture, and refrain from going further in their teaching. They believe in transmission of knowledge only; they do not permit knowledge to be acquired with objective, analytical, experimental methods. They do not bother to have intellectual discussions with their students, either because it is time or energy-consuming, or because discussions mean communication with students. Some lecturers do not want to communicate on an individual basis with their students.

I know some lecturers who do not even look in the eyes of their students. They just lecture and leave the classrooms. Their students are afraid of asking questions because they would be taking risks in doing so. They might get very angry. They wouldn't allow it anyway. They don't believe in discussions. I believe that very few lecturers renew themselves or their points of views. Very few of them try to follow modern teaching techniques. Maybe they are not even aware of them. This makes teaching mechanical, not creative. (Primary Education)

Some of the interviewees talked about the shortcomings of the teacher training programs in Turkey. They complained about how the Higher Education Council requires universities to apply to the government's package-programs. Following these ready-made programs leave no time for liberal arts courses.

I can only talk about my own faculty. Recently teaching programs become more like technician programs. They no longer contain courses like sociology of education or philosophy of education, which emphasize the intellectual, sophisticated sides of teaching. This is a big problem, you know. The students do not have the ... how do you say? ... the spirit of devotion any more. They see teaching just as a job like any other. These forgotten, ignored courses, actually, show the constructivist side of teaching, show the philosophy, the meaning of teaching.

So you say only specific courses are offered in all departments?

Yes, only specific courses. They have specific purposes: to give knowledge only in those fields. If I am going to be an English teacher I

take only courses from that department. I do not take courses from, for example, philosophy, geography, history, or any course from here or there. This has two drawbacks: The students are refrained from general culture, and more important they are refrained from getting experience from other departments. They are only socialized in their own departments; they only get their own department's mentality, or education. Students are prevented, in a way, from enlarging their horizons. (Educational Sciences)

In line with the above quotations, it is claimed in the relevant literature that some faculty members are "ignorant about the choice and use of pedagogical methods which are suitable for an education that relies more and more on higher level cognition and interpersonal abilities" (Cowan, 1998, p. 30). These faculty members are either too ignorant or too arrogant, as mentioned in the quotations, to admit to themselves that their techniques are old-fashioned. Apart from the basic knowledge that people should possess, the knowledge that faculty members convey to their students should change over time, according to the age they live in. "We enter a world in which knowledge doubles in less than five years (the projection is that by the year 2020, it will double every 73 days). It is no longer feasible to anticipate our future information requirements" (Beyer, 1997, p. x). This knowledge revolution requires "higher-order" work skills, including communications skills, problem solving, and reasoning - the "skills of the twenty-first century" (Grub & Lazerson, 2005, p. 1). So how can we serve our students in such an age? The only sensible way is to create learning environments in classrooms such that they learn where to find what they are looking for, where to find reliable knowledge, and how to differentiate it from the unreliable one - in short, to encourage them to acquire the habit of life-long learning.

There is another drawback of teaching only technical knowledge: it does not overlap with content and pedagogy. Mishra and Koehler (2006) presented technology as a separate and independent knowledge domain, in contrast to the other domains, such as pedagogy and content. They used Shulman's (1986, p. 9) idea of pedagogic content knowledge, which is the knowledge overlap in education between subject knowledge and knowledge of appropriate pedagogy. The argument is that if the faculty want to be effective, they need to know not only about the subject matter and how to teach, but understand appropriate pedagogies for that particular subject, topic, or concept (Unwin, 2007). Mishra and Koehler (2006) suggest that technology knowledge needs to overlap with content and pedagogy, in order to create a more desirable outcome in education.

The Identities Of Higher Education Institutions

In this category, the academicians agreed that universities should have an identity, a distinctive individual point of view. This identity does not occur immediately; it takes time to develop. It is not eternal, either; it changes over time, as the demographics of the faculty members and students change.

I see universities as living organisms. It is just like us; we all have identities so universities have identities because they contain humans. We may not see them with our eyes; there is nothing concrete. What kind of a university am I? What kind of students do I want to accept? What kind of teachers do I want to recruit? I believe in [a] code of ethics; it encompasses dignity, worth of a person, human relationships, integrity and such. Then education at my university will be accordingly. If I believe in something else, then the university will carry those characteristics. Therefore, if you look at a campus you can understand the identity of that university. But maybe in a couple of decades the identity will change as people change. Actually it should change as time goes by; it should adapt itself to modern concepts. (Educational Sciences)

The university climate influences its identity, which in turn, enables students to acquire various value systems (Özer & Demirtaş, 2010) and traditions. For example, critical thinking is perhaps the most important among the characteristics determining the multi-faceted academic identity of universities (Chowaniec, 2005). The elective courses students take and the extra curricular activities they choose on campus can be the second and third determinant factors. The atmosphere that is created in these environments, the relationship between the faculty members and students, students and students, or between students and their advisors, the faculties and the outside world all are of utmost importance in building the identity of the university, and, as a consequence, shaping the students' understanding of civic engagement and their capacity for social responsibility.

Becoming a lawyer who is fully equipped only with professional knowledge is not enough, in my opinion. A lawyer who has different hobbies, or interests, who spends her/his spare time in doing useful things other than their professional engagements... That is what is important. Universities have responsibilities in enabling their students in getting the habit of engaging in different hobbies, in engaging different useful activities. (Primary School Education)

The Roles of Faculty Members

In this category, the academicians all believed that the roles of faculty members are threefold: to teach, to do research, and to help students become active and conscientious. It was a common belief among the interviewees that the faculty members are not only transmitters of knowledge but organisers and facilitators of learning processes at the same time. They are the people who enable students to develop thinking and social abilities, which help produce cultural models.

How they teach is more important than what they teach because the way one teaches that particular knowledge makes all the difference. It is what stays in the minds of the students. The faculty should not only evaluate the level of knowledge of a student, they should also try to give importance to values in a society. Values! This is the glue that

holds a society together. If you examine who is responsible for most of the corruption in this country, you will see that they are all graduates of higher education institutions. Amazing, isn't it? (Foreign Language Education)

The other important point emphasized by interviewees was the relationship between the roles of the faculty members and the knowledge that is created in the classrooms. Their personalities and approaches to their students or to any kind of problem in the classroom, and certainly the way they treat and lecture their students are all factors affecting the knowledge that is shared in the classroom. It is the faculty's job to encourage their students how to think critically, to analyze, to deal with problem-solving techniques, and to work independently, as well as in a group. The faculty bring important insights, which constitute an important perspective on the nature of the institutional conditions necessary for learning (Nixon, 1996).

If the teacher stands in front of her/his students as an authoritarian figure, even if s/he does not tell straightforward but insinuates that s/he is the only person in the classroom possessing knowledge, this approach inhibits students to be active, to ask questions, to have discussions with each other. We have to see our classes as communities. I want them to join in the discussions, to reach knowledge by their own efforts, to inquire all together. I even want to learn from them. But when you become an authoritarian figure then they withdraw, they are not motivated any more, they become passive trying to get knowledge directly from you without thinking critically. (Primary School Education)

We know from recent literature that student-centered learning environments are more likely than teacher-centered learning environments to encourage students to adopt a deep approach to study (Gow & Kember, 1993; Kember, 1997). The techniques used – hands-on learning, learning through problem solving, group activities – need communication between teacher and students, and students and students. It is believed that in the traditional classrooms of HEIs, the faculty members take a variety of roles: cognitive roles, affective roles, disciplinary roles, managing roles, evaluative roles, performing roles, facilitator roles, gatekeeper roles, and boundary-spanner roles, which have one critical factor in common: verbal and nonverbal communication that the faculty members employ to enact these various roles (Coppola, Hiltz, & Rotter, 2002).

There are several studies discussing what it is like to be an academician. Although many lecturers consider themselves apart from schoolteachers and visualise themselves more as a member of their discipline (Becher, 1989), some think that being an academician is an art, and others think that it is a science. Those who view it as an art believe that academe encompasses aptitude, love, creativity, and inspiration; those who think it is science believe that it is knowledge and skill-based. The former is about personality characteristics, which are more difficult to influence than the latter, which deals with enabling people to get knowledge and skill.

Academics should be careful not only with getting knowledge and skill, but should work on their personality characteristics as well ("Akademisyen ve Roller," 2008). In an interview, one of the oldest and most prestigious academics of a state university describes her aim in teaching (Arman, 2008, p. 5):

I want my students to gain a sense of responsibility, curiosity, and excitement. And of course, honesty and open-mindedness. After graduating, I want them to have the motivation of contributing to the world. I want them to get used to keeping their promises. Doing what you have promised is important: not to be late, to answer a mail on time, to call someone back. I teach them to be good people - this is my ultimate aim in teaching.

Due to massification in HEIs, it is difficult to be a teacher and the burnout level among the faculty is high (Küçükşüleymanoğlu, 2007). Students' diversity relates not only to academic achievement, but also to ability, disability, age, sex, race, religion, class, maturity, experience, commitment, motivation, study mode, and the like. If higher education is to provide education for all, then the education provided should be suitable for all (Davis, 2003).

Conclusion and Implications for Practice

This study has identified the purposes, knowledge, identities, and roles of the faculty members of Turkish higher education institutions, as perceived by some of the practitioners. According to the results, we can say that the above factors are closely interrelated. Academics are founders and applicants of the purposes and identities of their universities, which in turn determine the identity of the universities. Academics also transfer or convey this identity to their students through the knowledge they transmit in their courses. Students who are educated with this knowledge and identity graduate from these universities and exist in their societies carrying the characteristics of the identity of their universities. This is a process that takes time; it does not happen immediately. Therefore, setting goals and/or purposes for a university may be the first step in higher education.

Ethical, social, and cultural values are of highest significance, but we also know that there is "an increasing demand on all the workforce to have better skills, and economic growth will depend on the skills, inventiveness and creativity of the workforce" (Blackstone, 2001, p. 176). She further says that higher education, which has a central role in the development of the knowledge economy, is worth 300 billion pounds per year, and by 2025, the estimated demand for higher education will be 159 million global enrollments with 87 million from Asia. What is more, there is also an expansion, which is called borderless education where web-based materials and programs can be delivered worldwide at the click of a mouse. Due to the changing political, social, and economic climate, with the growth of the private university sector, the traditional role of the public higher education sector is being questioned. Therefore, in order to remain viable actors in a changing political, social, and

economic environment, universities must redefine their roles in their societies (Twombly, 1997).

Without an understanding of the purposes of HEIs, concepts about what to learn are meaningless. Rowland (2002) claimed that although academicians are required to specify the aims and objectives of their teaching, the wider purposes go largely unexamined. The purpose of higher education is not only the development of people, or of society, or of economy, but people, society and the economy all together; furthermore, higher education, in turn, develops with these three themes in mind (Ranson, 1998).

The link between education and the corporate world is an important one, but it becomes more important when universities want to do research to an extensive degree. Although higher education budgets get larger and larger shares from the national budgets, most of them do not possess the amount of money to start or keep major research studies going. They need the outside world to back them up financially. Therefore, they have to get involved with national, international, global companies, institutions or foundations. It is difficult to keep research expenses low in the twenty-first century. Long gone are the days of the ancient Greek scholar, Archimedes, who experienced one of the greatest discoveries of all time at home in his bathtub, proclaiming "Eureka!" Studies now need millions and/or billions of U.S. dollars. Very few universities – if any – in the world have this amount of money available for research. Therefore, it is unavoidable for universities to be involved with the corporate world.

There is the other side of the coin, which is the need the corporate world has for the universities to produce research for them. Employers of private companies, organizations, and public enterprises need research increasingly in order to do their job properly. "Thus there is a two-way development of steadily stronger inter-relationships and mutual influences.... Although this may expose universities to a pressure to become more useful, this utilitarian pressure is not uniform because the needs of those who express them are more varied than ever" (Bleiklie & Byrkjeflot, 2002, p. 524).

Chickering (2002) claims that in order to foster critical thinking, multi-cultural understanding, and civic responsibility among students, HEIs can prepare students not only for a job but for a satisfying and productive career as well, which requires interpersonal competence, multi-cultural understanding, skills in problem identification and problem solving, a sense of purpose, and the confidence that you can act in ways that make a difference. "So if we only focus on specific professional and vocational preparation we sell both our students and our society short" (p. 7).

At this point, it can be said that one of the most salient duties in HEIs lies with the individual teacher, who wants to shape her/his students. Although there is an enormous amount of research on teachers and teaching, we know relatively little about the sources of influence, which shape the way they view their roles. It can be "their own personal and social background, their professional training, their policies and practices in the school in which they work, their access to collegial influence and

professional development and the individual's personality" (Osborn & Broadfoot, 1993, p. 105). In spite of these various reasons they have for their roles, we know from some studies that the main source of motivation for the faculty is their students (Kızıltepe, 2008). This gives us hope that, despite the changing world, and the changing higher education systems in the world toward more utilitarian characteristics, faculty members are aware of the caring role they have for their students. Walker, Gleaves and Grey (2006) claim that academicians manage considerable amounts of caring work with a diverse body of students who often need academic and pastoral support. Therefore, even if the conditions of their universities do not allow them to act on the institutional level, the faculty members individually can work wonders with their students, preparing them to be informed, active, and critical citizens of the twenty-first century.

Limitations and Suggestions for Future Research

The present study has a number of limitations. One of these arises from the use of the constant comparative method. Although it constitutes the core of qualitative analysis in different kinds of qualitative research, the application of the method remains rather unclear, and the analysis process needs to be systematized, and the traceability and verification of the analyses to be increased (Boeije, 2002).

Secondly, although the university in which this study took place is quite a large institution, it cannot be said that it is representative of the perspectives of the rest of the faculty working in different HEIs in different parts of Turkey. Therefore, this study has not attempted to generalize the findings and has not made claims relevant to all the faculty members in the country. It is limited to the perspectives of eighteen faculty members.

Finally, due to convenience sampling, the study has been conducted with female faculty members only. It would be interesting to see if male faculty members share their colleagues' ideas. In addition to that, looking into the comparison between male and female members could be worthwhile.

To sum up, future research needs to be done with faculty members in universities in all parts of Turkey, to see a more generalized picture. Furthermore, it needs to be done not only in the faculties of education but in other faculties as well, to compare and contrast the perspectives. This issue of purposes, identities of universities, the knowledge transmitted, and the role of the faculty members working in those institutions needs to be further researched in Turkey.

References

- Akademisyen ve rolleri (An academician and her/his roles). Retrieved June 12 2008 from <http://www.akademikkariyer.net/i.aspx?s=&k=haber&hab=m&i=23>.
- Akerlind, G. S. (2004). A new dimension to understanding university teaching, *Teaching in Higher Education*, 9(3), 364-375.
- Arman, A. (2008, February 3). Nihai amacım öğrencilerimi daha iyi insanlar yapmak (My ultimate aim is to make my students good people), *Hürriyet*, p. 5.
- Becher, T. (1989). *Academic tribes and territories: intellectual enquiry and the cultures of disciplines* (Milton Keynes, SRHE and Open University Press).
- Beyer, B. K. (1997). *Improving student thinking: a comprehensive approach* (Boston, Allyn and Bacon).
- Blackstone, T. (2001). Why learn? Higher education in a learning society, *Higher Education Quarterly*, 55(2), 175-184.
- Bleiklie, I. (2005). Organizing higher education in a knowledge society, *Higher Education*, 49, 31-59.
- Bleiklie, I., & Byrkjeflot, H. (2002). Changing knowledge regimes: universities in a new research environment, *Higher Education*, 44, 519-532.
- Boeije, H. (2002). A purposeful approach to the constant comparative method in the analysis of qualitative interviews, *Quality & Quantity*, 36, 391-409.
- Bryce, T. & Humes, W. (2000). *Scottish education*. Edinburgh: Edinburgh University Press.
- Chickering, A. W. (2002). Reclaiming our soul: democracy and higher education, *Change*, 35 (1), 39-45.
- Chomsky, N. (2000). Paths taken, tasks ahead, *Profession 2000*, 32-39.
- Chowaniec, J. (2005). Exploring identity, *Teaching in Higher Education*, 10(2), 265-270.
- Coppola, N. W., Hiltz, S. R., & Rotter, N. G. (2002). Becoming a virtual professor: Pedagogical roles and asynchronous learning networks, *Journal of Management Information Systems*, 18(4), 169-189.
- Cornwell, G. H., & Stoddard, E. W. (2001). The future of liberal education and the hegemony of market values, *Liberal Education*, 87(3), 6-16.
- Cowan, J. (1998). *On becoming an innovative university teacher: Reflection in action* (Milton Keynes, Open University Press).
- Çakmakçı, G. (2009). Preparing teachers as researchers: Evaluating the quality of research reports prepared by student teachers. *Eurasian Journal of Educational Research*, 35, 39-56.
- Davis, M. (2003). Barriers to reflective practice: The changing nature of higher education, *Active Learning in Higher Education*, 4(3), 243-255.

- Delanty, G. (2001). *Challenging knowledge: The university in the knowledge society* (Buckingham, Society for Research into Higher Education/Open University Press).
- Demircioğlu, O. (2007). Gazi Üniversitesi Merkez Kütüphanesi: Bilgiyi hayata dönüştüren kütüphane (Gazi University Central Library: The library that transforms information to life), *Bilgi Dünyası*, 8(1): 167-173.
- Eddy, E. D. (1957). *Colleges for our land and time: The land-grant idea in American education*. New York: Harper.
- Erkoç, M. (2010). Personal communication with the Director of Student Affairs, 9 February.
- Erzinçlioğlu, Z. (1999). British Universities: An ethical dilemma, *Contemporary Review*, 274, p. 143-147.
- Gibbons, M. (1998). Higher education relevance in the 21st century, paper presented at the UNESCO World Conference on Higher Education, 5-9 October, Paris, France.
- Glaser, B. G. & Strauss, A. L., (1967). *The discovery of Grounded Theory: Strategies for qualitative research* (Chicago, Aldine Publishing Company).
- Gow, L., & Kember, D. (1993). Conceptions of teaching and their relationship to student learning, *British Journal of Educational Psychology*, 63, 20-33.
- Grubb, W. N., & Lazerson, M. (2005). Vocationalism in higher education: The triumph of the education gospel, *The Journal of Higher Education*, 76 (1), 1-25.
- Gültekin, N., Çelik, A., & Nas, Z. (2008). Üniversitelerin kuruldukları kente katkıları (Contribution of universities to the cities they are founded in), *Electronic Journal of Social Sciences*, 7(24), 264-269.
- Kaya, Y. K. (1989). *İnsan yetiştirme düzenimize yeni bir bakış: Eğitimde model arayışı. (A new view to our system of raising people: Searching for a model in education)* (Ankara, Bilim Yayınları).
- Kember, D. (1997). A reconceptualisation of the research into university academics' conceptions of teaching, *Learning and Instruction*, 7, 255-275.
- Kempner, K., & Taylor, C. (1998). An alternative assessment to higher education outcomes: Differentiating by institutional type, *Higher Education*, 36, 301-321.
- Kızıltepe, Z. (2008). Motivations and demotivations of university teachers, *Teachers and Teaching: theory and practice*, 14(5-6), 515-530.
- Küçüksüleymanoğlu, R. (2007). Burnout syndrome levels of education faculty academic staff, *Eurasian Journal of Educational Research*, 28, 101-112.
- Kwiek, M. (2003). Academe in transition: Transformations in the Polish academic profession, *Higher Education*, 45, 455-476.
- Lankshear, C. & Knobel, M. (2006). *A handbook for teacher research: From design to*

implementation. New York: Open University Press.

- Lazerson, M. (1998). The disappointments of success: Higher education after World War II. *Annals of the American Academy of Political and Social Science*, 559, 64-76.
- Lutz, F. W., & Field, R. W. (1998). Business valuing in academia, *Higher Education*, 36, 383-419.
- McCollough, M. A., & Gremier, D. D. (1999). Guaranteeing student satisfaction: An exercise in treating students as customers, *Journal of Marketing Education*, 21(2), 118-130.
- Merriam, S. B. (1998) *Qualitative research and case study applications in education* (San Francisco, CA, Jossey-Bass Publishers).
- Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge, *Teachers College Record*, 108(6), 1017-1054.
- Nixon, J. (1996). Professional identity and the restructuring of higher education, *Studies in Higher Education*, 21(1), 5-16.
- Oğuz, A. (2004). Bilgi çağında yüksek öğretim programları (Higher education programs in the information age), *Milli Eğitim Dergisi*, 164 (2), 1-8.
- Osborn, M., & Broadfoot, P. (1993). Becoming and being a teacher: The influence of the national context, *European Journal of Education*, 28(1), 105-116.
- Özer, N., & Demirtaş, H. (2010). Students' perceptions regarding the fairness of learning environment in faculty of education. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 38, 126-145.
- Özgüngör, S. (2010). Identifying dimensions of students' ratings that best predict students' self-efficacy, course value, and satisfaction. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 38, 146-163.
- Ranson, S. (1998). *Inside the learning society* (London, Cassell).
- Rhoades, G., & Slaughter, S. (1997). Academic capitalism, managed professionals, and supply-side higher education, *Social Text*, 15(2), 9-38.
- Rowland, S. (2001). Teaching for democracy in higher education, paper presented at *Higher Education Close-Up* conference, University of Lancaster, UK, 16-18 July.
- Rowland, S. (2002). Overcoming fragmentation in professional life: The challenge for academic development, *Higher Education Quarterly*, 56(1), 52-64.
- Sherman, R. R., & Webb, R. B. (1988). Qualitative Research in Education: A focus. In R. R. Sherman & R. B. Webb (eds.), *Qualitative Research in Education: Focus and Methods*, 2-22. Bristol, P.: Falmer Press.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching, *Educational Researcher*, 15(2), 4-14.

- Sönmez, V. (2003). Yüksek öğretimin yeniden yapılanması üzerine bir deneme (On the reconstruction of the higher education system), *Eğitim Araştırmaları*, 12, p. 4.
- Tufan, O. (2008). Büyüyünce ne olacağım? (What shall I become when I grow up?). Retrieved July 24 2008 from http://www.dogrutercih.com/buyuyunce-ne-olacagim-469_.html.
- Twombly, S. B. (1997). Curricular reform and the changing social role of public higher education in Costa Rica, *Higher Education*, 33, 1-28.
- Unwin, A. (2007).The professionalism of the higher education teacher: What's ICT got to do with it?, *Teaching in Higher Education*, 12(3), 295-308.
- Walker, C., Gleaves, A., & Grey, J. (2006). A study of the difficulties of care and support in new university teachers' work, *Teachers and Teaching: theory and practice*, 12(3), 347-363.

Yüksek Öğretim Kurumlarının Amaçları, Kimlikleri ve Akademisyenlerin Rollerini

(Özet)

Problem Durumu: Yüksek öğretim kurumlarının amaçları, kimlikleri, bu kurumlarda iletilen bilgi ve akademisyenlerin rolleri son zamanlarda birçok araştırmacı tarafından incelenmektedir. Bilgi, kültür merkezli olmaktan çıktığı ve daha çok kullanılabilirlik merkezli bir hale geldiği için, yüksek öğretim kurumları sanki klasik anlamda işe yararlılığını yitirmiş gibi algılanmaktadır. Üniversitelerin amaçları hakkındaki çağdaş inançlar, onların yönlerini klasik müfredatlardan işlevsel olanlara doğru döndürmelerine sebep olmuştur. Üniversiteler, politikaları, programları ve uygulamalarıyla, öğrencileri daha iyi bir iş edinmeye ya da kamusal, bölgesel ve yerel ekonomik ihtiyaçlara hazırlamak için, artık kamu yararından çok sanki kişisel çıkarlara hazırlıyor gibi düşünülmeyle başlamıştır.

Araştırmanın Amacı: Bu makale, bir devlet üniversitesinin eğitim fakültesinde çalışan akademisyenlerin bakış açılarından, yüksek öğretim kurumlarının amaçları ve kimlikleri, bu kurumlarda üretilen ve iletilen bilgiyi ve bu konularla bağlantılı olarak akademisyenlerin rolleri hakkında yapılan bir çalışmanın sonuçlarını incelemektedir.

Araştırmanın Yöntemi: İstanbul'da bulunan bir devlet üniversitesinin eğitim fakültesinde çalışan onsekiz akademisyenle yarı yapılandırılmış bire bir mülakatlar gerçekleştirilmiştir. Bu mülakatlar önceden üç kişi üzerinde pilot olarak test edilmiştir. Akademisyenler sadece yazar tarafından görüşmeye tabi tutulmuş; her görüşme kırk ile seksenbeş dakika arasında bir sürede meydana gelmiştir. Görüşmeler kayda alınmış, kelimesi kelimesine yazıya dökülmüş ve yakından incelenmiştir. Data analizi yaparken, data kaynakları içinde sürekli tekrarlanan temaları tespit etmek için sabit karşılaştırma yöntemi uygulanmıştır.

Araştırmanın Bulguları: Çeşitlilik gösteren görüşler bulunduğu halde, denekler yüksek öğretim kurumlarının ana amaçlarının "aktif vatandaşlık" konusunu geliştirmek olduğu üzerinde hemfikir olmuşlardır. Bunun yanı sıra, üniversitelerde oluşturulan ve iletilen bilgi, öğrencileri işyerlerine hazırlamanın yanı sıra, insanlığın refahı ve esenliğine de katkıda bulunmalıdır. Özellikle belirtilen önemli bir nokta da, yüksek öğretim kurumlarının belirli bir kimliğinin olması gerektiği; bu kimliğin ise akademisyenler tarafından içselleştirilip öğrencilere aktarılması gerektiği olmuştur.

Araştırmanın Sonuçları ve Önerileri: Sonuç olarak önerilecek noktalar şunlardır: Yüksek öğretim kurumlarının amaçlarını tespit etmeden ne öğretileceği ile ilgili planlar yapmak anlamsızdır. Değişen politik, sosyal ve

ekonomik dünyada aktif oyuncular olarak kalmak için, üniversiteler toplum içindeki rollerini yeniden gözden geçirmelidirler. Yüksek öğretimin amacı, sadece kişileri ya da toplumu ya da ekonomiyi geliştirmek değil; kişilerin, toplumun ve ekonominin hepsini birden geliştirmek ve bu süreç içinde kendisinin de gelişmesidir. Bu, kişilerarası ilişkileri, çoklu kültür anlayışını, problem belirlemede ve çözümede beceriyi, bir amaç belirlemeyi ve fark yaratabilecek bir şekilde davranmayı gerektirir.

Anahtar Sözcükler: Güdülenme, şevk kırılması, akademisyen, üniversitenin amaçları, üniversitenin kimliği.

A Turkish Adaptation of the Inventory of Parent and Peer Attachment: The Reliability and Validity Study

Ercan Kocayörük*

Suggested Citation:

Kocayörük, E. (2010). A Turkish adaptation of the inventory of parent and peer attachment: The reliability and validity study. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 40, 133-151.

Abstract

Problem Statement: Adolescent attachment with their parents has been explicitly considered in many studies, and a vast number of studies provide empirical evidence for the link between parental attachment and psychological well-being and adjustment in adolescence. Adolescents explore close and supportive relationships outside of their immediate family in this period, and attachment behavior is often directed beyond parental figures. Attachment theory provides a theoretical underpinning about the link between parent and child attachment style with peer-group functions and relationships. Therefore, in the current study, psychometric properties of the Inventory of Parent and Peer Attachment were examined in a sample of subjects in middle adolescence (aged between 14 and 18 years).

Purpose of the Study: The study aimed to examine the psychometric properties of the Inventory of Parent and Peer Attachment developed by Armsden and Greenberg, Turkish version (IPPA-T), with a sample of subjects in middle adolescence (aged between 14 and 18 years).

Methods: In the study, the reliability and validity of a three-dimension model of adolescents' attachment to mother, father and peer were examined. In order to determine how well the identified model of the original version of the IPPA fits the Turkish adaptation of IPPA (IPPA-T), a Confirmatory Factor Analysis (CFA) and Exploratory Factor Analysis (EFA) were performed to understand the factor structure of IPPA-T. 315 high school students from Ankara, Turkey, participated in the study. Test-retest reliability was conducted using a Pearson product-moment

* Assist. Prof. Dr., Çanakkale Onsekiz Mart University, Faculty of Education, kocayoruk@comu.edu.tr

correlation coefficient over a two-week interval. Finally, the validity analysis was conducted by correlating the total and subscale of IPPA-T scores with Positive and Negative Affect Scale and Self-Esteem Scale.

Findings and Results: The results of the initial confirmatory factor analyses suggested that the original factor structure of the IPPA does not fit the data for adolescents' ratings of mothers, fathers and peer on the measures. The exploratory factor analyses revealed a new three-factor structure for IPPA with a shortened scale. The second confirmatory factor analyses demonstrated that the new three-factor model provides an acceptable fit.

Conclusions and Recommendations: It was concluded that the IPPA-T constitutes a useful tool for the assessment of both parent and peer attachment in adolescents aged between 14 and 18 years. The findings of the current study provide a good starting point for further attachment instruments in the Turkish culture.

Keywords: Adolescence, attachment, parents, peer, validation.

Bowlby's (1969) attachment theory has provided a theoretical point of view on the strong affective bond established between infants and the primary caregivers. The basic premise of attachment theory is that the quality of the attachment relationship stems from the interaction between infants and their caregivers, particularly the degree to which they can rely on attachment figures for security and support. Caregivers (especially mothers) who are sensitive and consistently responsive to their infant's needs are likely to foster secure attachment in their children. As a result of the early attachment experience with caregivers, individuals construct an internal working model of themselves, other and relationship that they use to guide their expectations in subsequent close relationships (Bowlby, 1988). Therefore, individuals' experience with the availability of attachment figures in their lives shape their feelings of security and trust in others.

Over the years, in line with Bowlby's framework, many researchers have conceptualized attachment as a life span construct (Bartholomew, 1993; Rice, 1990), and have claimed that children maintain attachments bonds to their parents across childhood and into adolescence. Reflecting the increased recognition of the importance of attachment across a person's life span, the research into the attachment relationship between parents and adolescents in a peer context is increasingly being conducted. Consistent with Bowlby's attachment theory, various instruments or self-report measures of attachment have been developed and carried out to assess attachment in adolescence. Perhaps the most cited instrument for attachment is the "Inventory of Parent and Peer Attachment" (IPPA), developed by Armsden and Greenberg (1987). IPPA is a self-report measure of attachment, and it measures psychological security derived from relationships with a subject's mother, father and close friends in three subscales (Communication, Trust and Alienation). It has been ascertained to have adequate internal consistency, test-retest reliability, and acceptable convergent validity in Western culture (Armsden & Greenberg, 1987; Gullone & Rabinson, 2005).

It is not surprising that methods and measures for studying attachment have attracted the attention of many researchers from diverse cultures (Raja, McGee, & Stanton, 1992; Vignoli & Mallet, 2004). Equally, it is not surprising that several researchers have tried to translate and adapt IPPA for the Turkish context, as previously there were no adequate, psychometrically sound and valid instruments for measuring the attachment of adolescents in Turkish culture. The first study using IPPA was conducted by Hortaçsu, Oral and Yasak-Gültekin (1991) to investigate the relationship between Turkish late adolescents and their parents and peers in a sample of undergraduate students between the ages of 16 and 20 ($M=20.26$ years). The item analysis and factor analysis were done with three separate varimax rotations, and the items were grouped by three factors, as Armsden and Greenberg (1987) pointed out in their study. Although the items of each subscale were not reported, the authors suggested that factor loadings were mostly similar for mother and father attachment items, and that three common subscales were obtained for fathers and mothers by selecting items with values greater than .30 on both analyses. Three additional scales were also constructed from attachment items for friends. The three scales, labeled "Trust," "Communication" and "Alienation," were constructed for subjects' mothers, fathers, and friends. The internal consistency coefficient for each subscale ranged from .68 to .90. However, the authors provided no data about test-retest reliability or other types of validity.

In addition to the study mentioned above, IPPA was employed to investigate the differential effect of parental and peer attachment on social and emotional loneliness among adolescents (Löker, 1999). A factor analysis with maximum likelihood revealed five factors accounting for 46 % variance; however, the factor solution in the current study did not fit the original scale factor solution. Although a high internal consistency estimate was found in the Löker study ($\alpha = .92$), the factors were not differentiated, the total score was thus used to measure both parental and peer attachment, and the reliability coefficient alpha for the two scales was .92. A further study was also carried out by Günaydın, Selçuk, Sümer and Uysal (2005) to examine the psychometric quality of IPPA using reliability and validity test techniques. The researchers employed the short form of the IPPA developed by Raja, McGee and Stanton (1992) to evaluate the attachment to mother and father. The results showed that three factors of IPPA (Communication, Trust, and Alienation) did not construct for the Turkish late adolescents, although the mother and father total scales each had high internal consistency and reliability, .88 and .90 respectively.

As a result, the previous adaptation studies of IPPA reported that established factor structure and loadings did not construct for the IPPA (Günaydın, Selçuk, Sümer, & Uysal, 2005; Löker, 1999) and provided no data about the test-retest reliability and any types of validity (Hortaçsu, Oral, & Yasak-Gültekin, 1991) in a sample of the Turkish late adolescents. However, the findings of the previous studies do not necessarily exclude the possibility that IPPA needs to be improved and enhanced to more effectively measure the adolescent attachment script to mother, father, and peers. A vast number of studies (Çolaklıoğlu & Gürçay, 2007; Eren Gümüş, 2010) were carried out to examine the scales to develop valid and reliable measurements for Turkish adolescents, because of the lack of instruments specifically developed and validated for Turkish adolescents. Therefore, the present study

suggests comprehensive statistical approaches to examine the psychometric properties of IPPA.

The first aim of the present study was to explore the psychometric properties of IPPA adapted for Turkish middle-adolescents (IPPA-T) by using comprehensive statistical procedures and providing data about the reliability and validity of IPPA-T. In addition, many studies (e.g. Lieberman, Doyle, & Markiewicz, 1999; Wilkinson, 2004) suggested that the attachment behavior of adolescents could not be assessed by a single dimension, and that the father and mother attachment behavior system for adolescents should be considered separately. This suggestion was the second aim of the present study, which was to examine the psychometric properties of attachment to mother (MA), father (FA) and close friends (PA) in a sample of middle adolescence by using the revised version of IPPA (25 items for each scale). In order to determine the psychometric properties of the IPPA, the exploratory and confirmatory factor analyses were carried out and the factor structure and internal consistency of IPPA were analyzed. Next, the test-retest reliability was conducted using Pearson product moment correlation. Finally, the correlations between the IPPA and self-esteem (Rosenberg Self Esteem Scale-RSS) and psychological well-being (Positive and Negative Affective Scale-PANAS) were analyzed to determine the convergent validity of the IPPA.

Method

Research Design

Initially, in order to determine how well the identified model of the original version of the IPPA would fit the Turkish adaptation of IPPA (IPPA-T), a Confirmatory Factor Analysis (CFA) was performed on the variance-covariance matrix using Lisrel 8.3 (Joreskog & Sörbom, 2003). CFA allowed for determining whether the factor structure of the Turkish version of IPPA was comparable to the factor structure of the original version of the IPPA (Zero Model, M_0). If the original factor structure did not fit the data for the IPPA-T, an Exploratory Factor Analysis (EFA) with principal component analysis would be performed to understand the factor structure of the IPPA-T. After that, based on the results of EFA, the second CFA would be carried out to determine whether the IPPA-T new factor structure would fit the data (Alternative model, M_1). For this analysis, the goodness-of-fit statistics were tested with χ^2 (a non-significant value that corresponds to an acceptable fit). Because χ^2 are known to increase with sample size and degree of freedom, the use of four indices is commonly suggested; (a) Standardized Root Mean Square Residual (S-RMR) < .08, (b) Root Mean Square Residual of Approximation (RMSEA) < .06, (c) Goodness-of-Fit Index (GFI) > .90, (d) Adjusted Goodness-of-Fit Index (AGFI) > .85, and e) Comparative Fit Index (CFI) > .90 (Hu & Bentler, 1999; Şimşek, 2006). In addition, Expected Cross-Validation Index (ECVI) and Model AIC, assumed to be more acceptable for lower fits (Joreskog & Sörbom, 1993), were also reported to allow comparison of zero with the alternative measurement model.

Participants

399 high school students from Ankara, Turkey, were asked to participate in the study during the 2007-2008 academic years. 84 participants were excluded from the study because of their incomplete answers. The results were analyzed for the remaining 315 participants (147 male, 168 female) aged between 14 and 18 years ($M=16.06$, $SD=1.02$).

Research Instruments

Inventary of parent and peer attachment (IPPA; Armsden & Greenberg, 1987). IPPA derives from the theoretical assumption of attachment theory, and assesses the positive and negative dimensions of adolescents' relationships with their parents and close friends. Specifically, the IPPA was developed in order to assess adolescents' perceptions of the positive and negative affective/cognitive dimension of relationships with their parents and close friends -- particularly how well these figures serve as sources of psychological security. The items in the original version of the IPPA, in a sample of college students aged between 16 and 20, demonstrated good internal consistency and were clustered into three factors through principal components analysis, namely "Communication", "Trust", and "Alienation" (Armsden & Greenberg, 1987).

All three scales (father, mother, and peers) consisted of 25 items and were translated into Turkish prior to administering the study. Two bilingual Turkish scholars independently translated each item, and compared their translations to resolve any disagreements. From this translation, a Turkish-English bilingual supervisor translated it back into English. The discrepancies emerging from this back-translation were discussed, and the adjustments to the Turkish translation of the IPPA were made.

Rosenberg self-esteem scale (RSS; Rosenberg, 1965). RSS is a ten-item self-report scale developed by Rosenberg (1965) for the purpose of measuring adolescent's global self-esteem. The scale has ten items formulated to assess the appraisal of individual. Five of the items are phrased positively, and other five items are negatively. The Turkish adaptation of RSS was established by Çuhadaroğlu (1986) through psychometric interviews with ninth, tenth and eleventh grade secondary school students, and the results revealed strong reliability validity properties in the middle adolescence sample.

Positive and negative affect scales (PANAS; Watson, Clark, & Tellegen, 1988). PANAS, developed by Watson, Clark, and Tellegen (1988), is a ten-item brief measure of affective evaluation of life with two dimensions. Positive affect (PA) reflects the extent to which a person feels enthusiastic, active, and alert. In contrast, "Negative Affect (NA) is a general dimension of subjective distress and unpleasurable engagement that subsumes a variety of aversive mood states, including anger, contempt, disgust, guilt, fear, and nervousness, with low NA being a state of calmness and serenity. The adaptation of the scale to Turkish was made by Gençöz (2000). Consistent with the original study, the result of the factor analysis revealed two factors accounting for the 44% of the total variance.

Findings and Results

Factor Structure

Initial Confirmatory Factor Analysis (CFA) with the maximum likelihood method was performed on the variance-covariance matrix in order to test whether the original version of the IPPA-T fitted the data. In the first model (Zero Model, M_0), it was assumed that all items of IPPA-T were constructed in the same factors as in the original version of IPPA. The results revealed that the original factor structure of the IPPA-T indicated a poor fit of the data. The RMSEA and S-RMR were higher than the cutoff value for an acceptable level, and the GFI, AGFI and CFI were below .90, indicating a poor fit of the zero model (M_0). Indeed, the results suggested that original factor structures of the IPPA-T showed poor goodness-of-fit-statistics for mother, father, and close friends or peers (Table 1).

Table 1

Fit Indices of Confirmatory Factor Analyses

Models	χ^2	df	SRMR	RMSEA	GFI	AGFI	CFI	ECVI	Model AIC
Mother									
Zero Model (M_0)	628.79*	272	.064	.065	.86	.84	.89	2.34	734
Alternative Model	282.86*	132	.055	.060	.91	.88	.94	1.15	360
Father									
Zero Model	730.12*	272	.088	.073	.84	.81	.89	2.66	836
Alternative Model	335.88*	132	.072	.070	.89	.86	.93	1.09	412
Peer									
Zero Model	743.11*	272	.070	.074	.84	.81	.87	2.70	849
Alternative Model (M_1)	275.07*	132	.053	.059	.91	.89	.93	1.12	353

* $p < .01$

Exploratory Factor Analysis: Because of a poor fit of the zero model (M_0), a principal components factor analysis with an oblimin rotation was conducted to further understand the factor structure of the IPPA-T. The factor solution was determined using the scree plot method. In addition, it was decided that the factors extracted would only comprise items with a factor load greater than .35.

For the mother attachment (MA) scale, the results showed that the factor solution for the 19 items were grouped into three factors, and 6 items (6, 9, 10, 13, 17, 23) were eliminated due to their low factor load (<.35). For the father attachment (FA) scale, a factor analysis was computed to the oblimin rotation. The eigenvalues were greater than one, suggesting that the 24 items were grouped into three factors. One item (4) was eliminated because it displayed a low factor load (<.35). On the peer attachment (PA) scale, the factor solution was mostly similar to the father attachment items, and the factor solution composed of 24 items was grouped into three factors. One item (6) was eliminated as it displayed a low factor loading (<.35).

In order to achieve a parallel form between the mother, father and peer attachment scales, and thus to provide an assessment of the relationship between father, mother, and peers, some items in each scale, which had a factor loading lower than .45 on the expected factors, were eliminated. Therefore, one item (4) from the mother dimension, six items (6, 9, 10, 13, 17, 18) from the father dimension, and six items (2, 5, 7, 8, 9, 21) from the peer dimension were excluded. Consequently, 18 items for attachment to mother, father and peers were determined, and three separate factor analyses were performed for each (Table 2).

Table 2

Loadings of the Exploratory Factor Analysis

Items	Factor 1	Factor 2	Factor 3
<i>Mother</i>	1	2	3
19 My mother helps me to talk about my difficulties.	.85	-.25	.05
20 My mother understands me.	.81	-.34	.27
12 When we discuss things, my mother cares about my point of view.	.79	-.07	.27
1 My mother respects my feelings.	.77	-.13	.27
24 I can count on my mother when I need to get something off my chest.	.76	-.30	.14
15 My mother helps me to understand myself better.	.75	-.30	.08
21 When I am angry about something, my mother tries to be understanding.	.72	-.24	.11
5 I like to get my mother's point of view on things I'm concerned about.	.63	-.16	.18
16 I tell my mother about my problems and troubles.	.62	-.51	-.01
25 If my mother knows something is bothering me, she asks me about it.	.60	-.12	.24

7	My mother can tell when I'm upset about something.	.50	-.11	.33
2	I feel my mother does a good job as my mother.	.28	.66	-.18
3	I wish I had a different mother.	-.12	-.66	.10
22	I trust my mother.	.30	.53	-.05
14	My mother has her own problems, so I don't bother her with mine.	-.08	-.23	-.76
11	I get upset a lot more than my mother knows about.	-.21	-.15	-.67
8	Talking over my problems with my mother makes me feel ashamed or foolish	-.27	.16	-.59
18	I don't get much attention from my mother.	-.10	-.25	-.58
<i>Father</i>				
19	My father helps me to talk about my difficulties	.87	.14	-.23
25	If my father knows something is bothering me, he asks me about it.	.85	.23	.07
15	My father helps me to understand myself better.	.82	.27	.17
7	My father can tell when I'm upset about something.	.77	.28	.13
21	When I am angry about something, my father tries to be understanding	.74	.10	.31
1	My father respects my feelings.	.74	.37	.18
24	I can count on my father when I need to get something off my chest.	.73	.07	-.34
16	I tell my father about my problems and troubles	.72	.26	.13
20	My father understands me.	.71	.42	-.26
5	I like to get my father's point of view on things I'm concerned about.	.61	.41	.08
12	When we discuss things, my father cares about my point of view.	.60	.42	.19
2	I feel my father does a good job as my father	.30	.78	-.03
3	I wish I had a different father.	.12	-.70	.09
22	I trust my father.	.28	.60	-.01
8	Talking over my problems with my father makes me feel ashamed or foolish	-.24	.07	-.71
14	My father has his own problems, so I don't bother him with mine.	.09	.10	-.66
11	I get upset a lot more than my father knows about.	.09	-.48	-.60
23	My father doesn't understand what I'm going through these days.	-.24	-.34	-.56
<i>Peers</i>				
17	My friends care about how I am feeling.	.78	.30	.22
3	When we discuss things, my friends care about my point of view.	.73	.23	.18

16	My friends help me to understand myself better.	.72	.41	-.01
15	When I am angry about something, my friends try to be understanding.	.71	.26	.19
14	My friends are fairly easy to talk to.	.68	.12	.26
20	I trust my friends.	.18	.83	.36
19	I can count on my friends when I need to get something off my chest.	.13	.80	.26
24	I can tell my friends about my problems and troubles.	.34	.76	.22
25	If my friends know something is bothering me, they ask me about it.	.35	.73	.14
1	I like to get my friend's point of view on things I'm concerned about.	.23	.70	.08
13	I feel my friends are good friends.	.20	.70	.45
12	My friends listen to what I have to say.	.34	.56	.11
22	I get upset a lot more than my friends know about.	-.01	-.20	-.77
18	I feel angry with my friends.	.02	.19	-.67
23	It seems as if my friends are irritated with me for no reason.	-.03	-.31	-.66
10	My friends don't understand what I'm going through these days.	-.28	-.02	-.65
11	I feel alone or apart when I am with my friends.	-.15	.03	-.61
4	Talking over my problems with friends makes me feel ashamed or foolish.	-.09	-.21	-.45

In the mother attachment, a three factor solution accounting for 54.50 % of the total variance was found. The first factor eigenvalue of the principal component analysis was 7.43 and accounted for 41.80 % of the variance and included eleven items (1, 5, 7, 12, 15, 16, 19, 20, 21, 24, 25). The second factor was grouped by the three items (2, 3, 22) and the eigenvalue of the principal component analysis was 1.20, which accounted for 6.07 % of the variance. The third factor eigenvalue of the principal component analysis was 1.28, accounted for 6.63 % of the variance and included four items (8, 11, 14, 18).

For the father attachment (FA) scale, the eigenvalues were greater than one, suggesting that the 18 items grouped themselves into three factors (eigenvalues for factor1= 8.20, factor2= 1.31, factor3= 1.40) with each factor accounting for a comparable amount of variance (45.56 %, 7.25 %, 7.74 %, respectively, in total accounting for 60.57 % of the variance). The first factor was characterized by eleven items (1, 5, 7, 12, 15, 16, 19, 20, 21, 24, 25). The second factor comprised three items (2, 3, 22) and the third factor included four items (8, 11, 14, 23).

In peer attachment, the factor solution was composed of 18 items constructed in three factors that accounted for 51.77 % of the variance. The first factor (eigenvalue= 1.45) accounted for 6.39 % of the variance and included five (3, 14, 15, 16, 17) items. The second factor (eigenvalue=8.50) was grouped by the seven items (1, 12, 13, 19, 20, 24, 25) and accounted for 36.13% of the variance. The last and the third factor

(eigenvalue=1.67) was composed of six items (4, 10, 11, 18, 22, 23) and accounted for 9.24 % of the variance.

The results of the principal components factor analysis revealed that three factor structures relating to adolescents' attachment to mother, father and peers were found. In other words, the oblimin rotation results pointed out that a three factor structure (Communication, Trust and Alienation) was obtained again for mother, father and peer attachment as shown in Armsden and Greenberg's study. The exploratory factor analysis results also showed that a shortened version of the IPPA was more appropriate for middle adolescents in Turkey.

Furthermore, the process of deciding the number of factors was carried out with a parallel analysis, a commonly used method that compares the size of the eigenvalues with those produced by a randomly generated data set (Horn, 1965). The results of parallel analysis also supported the retention of three factors in three dimensions (father, mother and peer). The eigenvalues of all of the three factors in each three dimensions exceeded the criterion values produced from the random data matrix of the same size (315X18 items) while the fourth factor, with an eigenvalue of 1.01, accounted for 5% of the variance, which reflected in an elbow formed at the fourth factor in the peer dimension. The further details of the eigenvalues generated from the principal axis and the criterion values obtained from parallel analysis are demonstrated in Table 3.

Table 3

A comparison of Eigenvalues from the Principal Axis factor Analysis and the Corresponding Criterion Values Obtained from the Parallel Analysis

Mother				
Component number	Actual eigenvalue	Criterion Value from PA	Value	Decision
1	7.43	1.40		Accept
2	1.20	1.007		Accept
3	1.28	1.26		Accept
Father				
Component number	Actual eigenvalue	Criterion Value from PA	Value	Decision
1	8.20	1.42		Accept
2	1.31	1.28		Accept
3	1.40	1.34		Accept
Peer				

Component number	Actual eigenvalue	Criterion Value from PA	Decision
1	1.45	1.34	Accept
2	8.50	1.45	Accept
3	1.67	1.43	Accept
4	1.014	1.285	Reject

Confirmatory Factor Analysis: In line with the statistical analysis strategy, a second confirmatory factor analysis (Alternative Model, M_1) was conducted to determine whether the new factor structure, based on the results of exploratory factor analysis, would fit the data for the population in this study. The results showed that the χ^2 value of the Alternative Model (M_1) was significant for mother attachment (MA) [$\chi^2_{(132)}=282.86, p<.01$], father attachment [$\chi^2_{(132)}=335.88, p<.01$], and peer attachment scales [$\chi^2_{(132)}=275.07, p<.01$]. However, other goodness-of-fit statistics of the Alternative Model (M_1) for each scale indicated a better fit to the data for IPPA-T than the Zero Models (M_0). The χ^2 value/degree of freedom ratio for the Alternative Model was slightly above 2, but this ratio was considered as acceptable by Jöreskog and Sörbom (1993). For other fit indices, better goodness-of-fit statistics were found for the Alternative Model (M_1) with more acceptable values for GFI (>.90), AGFI (>.85), S-RMR (<.080), CFI (>.90) and RMSEA (<.060) in mother, father and peer attachment scales, except RMSEA (>.060) and GFI (<.90) in father attachment. Furthermore, the values of ECVI and Model AIC were more acceptable for the alternative model that was lower than the zero model for adolescents' attachment to mother, father and peers (Table 1).

In addition, in order to test whether the alternative model would be replicated in an independent sample, the alternative model was examined in an independent sample ($N=226, M=15.35; SD=.85$). In the mother dimension, the results of goodness of fit statistics were: $\chi^2_{(132)}=274.28, p<.01$, S-RMR=.056, RMSEA=.069, CFI=.91, AGFI=.85, and GFI=.88. In the father dimension, goodness of fit statistics were: $\chi^2_{(132)}=318.17, p<.01$, S-RMR=.047, RMSEA=.079, CFI=.92, GFI=.86, and AGFI=.82. Lastly in the peer dimension, goodness of fit statistics were: $\chi^2_{(132)}=282.44, p<.01$; S-RMR=.055, RMSEA=.071, CFI=.92, GFI=.88, and AGFI=.84. The results suggested that fit indices were also acceptable in an independent sample.

Reliability

Internal Consistency: The Cronbach' Alpha (α) internal consistencies were .91 for total mother attachment, .92 for Communication, .63 for Trust, and .62 for Alienation. The Cronbach' Alpha (α) internal consistencies were .92 for total father attachment, .93 for Communication, .69. for Trust, and .66 for Alienation scale. The last Cronbach' Alpha (α) internal consistencies were .89 for total peer attachment, .80 for Communication, .85 for Trust, and .71 for the Alienation scale.

The results also revealed that each of the total IPPA-T scores highly correlated with their respective sub-scale scores. The total attachment to mother score was

correlated with .97, .62, and .73 for Communication, Trust, and Alienation, respectively. The total attachment to father was correlated with .97, .81, and .72 for Communication, Trust, and Alienation, respectively. The total attachment to peers was correlated with .84, .88, and .78 for Communication, Trust, and Alienation, respectively. Furthermore, as expected, the intercorrelations between subscales, within the scale (mother, father, and peer), were moderately high. Communication and Trust subscales were positively correlated with $r=.54$, $.77$, and $.67$ for mother, father and peer scales, respectively. In addition, the Alienation subscale was negatively correlated with Communication ($r=-.56$) and Trust ($r=-.33$) for attachment to mother, with $r=-.56$ and $-.46$ for father, and $r=-.47$ and $-.50$ for peers, respectively.

Test-Retest Reliability: Test-retest reliability was conducted using Pearson product-moment correlation coefficient over a two-week interval. The sample of the test-retest reliability study consisted of 77 adolescents ($M=15.35$, $SD=.85$) aged between 14 and 17. The result of the Pearson's product-moment correlation coefficient revealed that test-retest reliability was high in the mother and father subscale, but moderate in the peer subscale. In the mother subscale, Person's product-moment correlation was .73 for total mother attachment (MA), .72 for Communication, .34 for Trust, and .50 for Alienation. In the father subscale, Person's product-moment correlation was .81 for total father attachment (FA), .81 for Communication, .65 for Trust, and .63 for Alienation. Lastly, in the peer subscale, test-retest reliability was .55 for total peer attachment (PA), .51 for Communication, .53 for Trust and .46 for Alienation.

Validity

Convergent Validity with Self-Esteem and Well-Being: Validity analysis was conducted by correlating the total and subscale of IPPA-T scores with Positive and Negative Affect Scale (PANAS) and Rosenberg Self-Esteem Scale (RSS). The convergent validity method was used to determine whether the total score, Communication and Trust subscale scores of IPPA-T correlated significantly and positively with the RSS and PANAS positive well-being (PA) subscale score. In addition, a significant correlation was expected negatively with the PANAS negative well-being (NA) subscale scores. Conversely, it was expected that Alienation subscale scores correlated negatively with RSS and PANAS positive well-being (PA) subscale scores, but positively with the PANAS negative well-being (NA) subscale scores. The instruments were completed over four weeks and the voluntary nature of the adolescents' participation ($N=282$, $M=16.01$, $SD=1.04$) was clearly stated prior to administering the instruments.

The findings revealed that the total scores and Communication and Trust subscale scores of mother attachment and father attachment were positively correlated to the positive affect (PA) and RSS. In addition, the total scores of mother attachment and father attachment, and Communication and Trust subscale scores, were negatively correlated to the negative affect (NA). As expected, the Alienation subscale of two attachments (father and mother) was negatively correlated to PA and RSS, and positively correlated to NA.

In the peer attachment scale, the total score of peer attachment, Communication and Trust subscale scores were positively correlated to PA and RSS. Contrary to the expectations, while the total score of peer attachment negatively correlated to NA,

the Communication and Trust subscale did not correlate to NA. The Alienation subscale was negatively correlated to PA and RSS, and positively correlated to NA, as expected (Table 4).

Table 4

Correlations of the IPPA-T with PANAS and SE Measures

	Positive Affect	Negative Affect	Self-Esteem
Mother (Total score)	.39**	-.36**	.45**
Communication	.38**	-.32**	.40**
Trust	.36**	-.25**	.30**
Alienation	-.20**	.35**	-.42**
Father (Total score)	.36**	-.37**	.40**
Communication	.34**	-.30**	.34**
Trust	.30**	-.36**	.30**
Alienation	-.26**	.40**	-.40**
Peer (Total score)	.25**	-.13*	.27**
Communication	.22**	-.06 ^a	.22**
Trust	.19**	-.03 ^a	.16**
Alienation	-.23**	.27**	-.35**

** Correlations are significant at $p < .01$, * Correlations are significant at $p < .05$,

^a Correlations are insignificant

Conclusions and Recommendations

The results of the EFA and CFA revealed that the three factor structures (Trust, Communication, and Alienation), with 18 items for mother, father and peer attachment scales, were most likely replicated with Turkish middle age adolescents. The results also showed that the acceptable level of internal consistency was obtained for the new three-factor structure (18 items) and the total score of the IPPA-T. Test-retest reliability was moderately strong among the subscales and total score of the IPPA-T, and the results consistently agreed with the theoretical framework with respect to convergent validity.

The overall findings of the present study were in line with other adaptation studies of IPPA into diverse cultures. For instance, Vignoli and Mallet (2004), in France, found a clear and expected three-dimension model (Communication, Trust, and Alienation) for attachment to father, whereas no meaningful internal structure for parental attachment to mother was found. After the CFA and EFA studies, the authors applied a brief questionnaire for attachment to the mother with a three-dimension model (14 items) which could be considered as a reliable and valid measure of parental attachment for French late adolescents.

In addition, the results of the present study showed that the IPPA-T was a sound tool for the assessment of attachment in middle adolescents aged from 14 to 18, and this finding is in line with the studies that supported the validity of the IPPA construct for adolescent attachment to father and mother involving different age groups of adolescents. For instance, Gallone and Robinson (2004) conducted a study with different age groups (children and adolescents) to examine the revised version of IPPA. The authors found out that the revised IPPA was a sound tool for the assessment of attachment in children and adolescents aged from 9 to 15.

To summarize, although some of the items on the original scale were reduced, IPPA-T could be assumed to assess the same multidimensional constructs in the attachment relationship to father, mother and peers as proposed by Armsden and Greenberg (1987), and IPPA-T is a sound tool for the assessment of attachment in middle adolescents of ages 14 to 18 in the Turkish culture. The findings of the current study provided a good starting point for further IPPA development in the Turkish culture. In other words, further studies might try to devise and revise several new items to increase internal consistency of the IPPA-T, particularly the trust subscale, which consisted of only three items.

These findings should be considered within the context of the study's limitations. First of all, the sample of the current study was limited to middle adolescents, and the results may not be generalizable to samples of different ages. The second limitation of the study is that the small sample size was conducted using convergent validity analysis with PA, NA and RSS. This finding can be considered tentative and needs to be replicated. Another limitation of the study concerns reliability; meaning that the two-week measure period was too short to evaluate test-retest reliability, and evaluation would have needed to be extended over longer time intervals.

References

- Armsden, G.C. & Greenberg, M.T. (1987). The inventory of parent and peer attachment: Individual differences and their relationship to psychological well-being in adolescence. *Journal of Youth and Adolescence*, 16, 427-454.
- Bartholomew, K. (1993). From childhood to adult relationships: Attachment theory and research. In S. Duck (Ed.), *Learning about relationships* (pp. 30-62). Newbury Park, CA: Sage Publication.
- Bowlby, J. (1969). *Attachment*. Vol. 1 of Attachment and Loss. New York: Basic Books.
- Bowlby, J. (1988). *A secure base: Parent-child attachment and healthy human development*. New York: Basic Book.
- Çolakakadiođlu, O. & Gürçay, S., S. (2007). The Adaptation of Adolescent Decision Making Questionnaire into Turkish Population. *Eđitim Arařtırmaları- Eurasian Journal of Educational Research*, 26, 61-71.
- Çuhadarođlu, F. (1986). *Adolesanlarda benlik saygısı* [Self-esteem of adolescents]. Unpublished doctoral thesis. Hacettepe University. Ankara.
- Eren Gümüş, A. (2010). The construct validity, reliability of Self-Perception Profile for Adolescents: Original versus revised version. *Eđitim Arařtırmaları- Eurasian Journal of educational Research*. 39,127-144.
- Gençöz, T. (2000). Pozitif ve negatif duygu ölçeđi: Geçerlik ve güvenilirlik çalışması. [Positive and Negative Affect Scale: Reliability and Validity Study]. *Türk Psikoloji Dergisi*, [Turkish Journal of Psychology], 15 (46), 19-26.
- Gullone, E. & Robinson, K. (2005). The Inventory of Parent and Peer Attachment– Revised (IPPA-R) for Children: A Psychometric Investigation. *Clinical Psychology and Psychotherapy*. 12, 67-79.
- Günaydın, G., Selçuk, E., Sümer, N., & Uysal, A. (2005). Ebeveyn ve Arkadařlara Bađlanma Envanteri kısa formunun psikometrik açıdan deđerlendirilmesi [The psychometric evaluation of the short form of Inventory of Parent and Peer Attachment]. *Türk Psikoloji Yazıları*, [Turkish Psychology Articles], 8(16), 13-23.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Hortaçsu, N., Oral, A., & Yasak-Gültekin, Y. (1991). Factors affecting relationship of Turkish adolescent with parents and same-sex friends. *The Journal of Social Psychology*, 131, 413-426.
- Hu, L. T. & Bentler, P.M. (1999). Cutoff criteria for fit indexes incovariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jöreskog, K. G. & Sörbom, D. (1993). *LISREL 8. User's reference guide*. Chicago, IL: Scientific Software.

- Lieberman, M., Doyle, A., & Markiewicz, D. (1999). Developmental patterns in security of attachment to mother and father in late childhood and early adolescence: Associations with peer relations. *Child Development, 70*, 202-213.
- Löker, Ö. (1999) *Different effects of parent and peer attachment on social and emotional loneliness among adolescents*. Unpublished doctoral thesis. Middle East Technical University. Ankara.
- Raja, S. N., McGee, R., & Stanton, W. R. (1992). Perceived Attachments to Parents and Peers and Psychological Well-Being in Adolescence. *Journal of Youth and Adolescence, 21*, 471-485.
- Rice, K. G. (1990). Attachment in adolescence: A narrative and meta-analytic review. *Journal of Youth and Adolescence, 19*(5), 511-538.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Şimşek, Ö., F. (2006). *Yapısal eşitlik modellemesine giriş* [Introduction to Structural Equation Modeling]. Ankara: Ekinoks Yayıncılık.
- Vignoli, E. & Mallet, P. (2004). Validation of a brief measure of adolescents' parent attachment based on Armsden and Greenberg's three-dimension model. *Revue Européenne de Psychologie Appliquée* [European Review of Applied Psychology], *54*, 251-260.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.
- Wilkinson, R. B. (2004). The role of parental and peer attachment in the psychological health and self-esteem of adolescents. *Journal of Youth and Adolescence, 33*(6), 479-493.

Anne-Baba ve Akran Bağlanması Envanterinin Türkçe Uyarlaması: Güvenirlilik ve Geçerlik Çalışması

(Özet)

Problem Durumu: Bowlby yaşamın ilk yıllarında annenin/bakıcının çocuğa verdiği tepkilere bağlı olarak, çocuğun kendisine ve başkalarına ilişkin zihinsel modeller oluşturduğunu ve bu zihinsel modellerin daha sonraki yıllarda, yakın kişiler arası ilişkiler için bir rehber ve model görevi gördüğünü belirtmektedir.

Yaşamın ilk aylarından itibaren bebek ile temel bakıcı (anne) arasında gelişen bağlanma örüntüsünün, hem çocuklukta hem de yetişkinliğe geçişte bireyin ruh sağlığı ile ilişkisini ortaya koyan araştırmalar dikkat çekmektedir. Son yıllarda ergen ve yetişkin ilişkilerinde bağlanmanın rolünü inceleyen araştırma bulguları; yaşamın ilk yıllarında annenin çocuğa verdiği tepkilere bağlı olarak çocuğun kendisine ve başkalarına ilişkin modeller oluşturduğunu, daha sonraki yıllarda da bu modellerin yakın kişiler arası ilişkiler için rehber görevi gördüğünü ortaya koymaktadır. Bağlanmanın yaşam boyu devam eden bir süreç olarak değerlendirilmesiyle birlikte anne-babanın ve akran gruplarına bağlanmanın ergenler üzerindeki etkilerine yönelik çalışmalarda son yıllarda artış gözlenmektedir. Bu ilginin bir sonucu olarak Bowlby'nin görüşleri ile tutarlı olarak, ergenlik döneminde bağlanma sürecinin değerlendirilmesi amacıyla bir çok ölçme aracı geliştirilmiştir. Bağlanma kuramı ile çalışmalarını sürdüren araştırmacıların bu ölçme araçları içerisinde belki de en çok ele aldıkları ölçme aracı Armsden ve Greenberg tarafından geliştirilen Anne-Baba ve Akran Bağlanması Envanteri'dir (ABABE). Ölçme aracının bir çok farklı dile ve kültüre uyarlandığı görülmektedir. ABABE'nin Türk kültürüne uyarlama çalışmalarının gerçekleştirilmiş olması da şaşırtıcı değildir. Yapılan çalışmalar geç ergenlik dönemi (late adolescence) örneklemleri üzerinde gerçekleştirilmiştir. Bu çalışmalardan bazılarında Armsden ve Greenberg'in belirttiği faktör yapılarının ayrıştığı görülmüş ancak çalışma sonuçlarında test-tekrar test ve geçerlilik çalışmaları rapor edilmemiştir. Diğer çalışmalarda ise faktör yapılarının ayrışmadığı sonucu bildirilmiştir. Genel olarak bakıldığında, uyarlama çalışmaları sonucunda ABABE'nin Türkiye'de bağlanma alan yazınında ve araştırmalarında kullanılabilir bir ölçme aracı olabileceği düşünülmektedir. Bu çalışmada ise ABABE'nin orta-ergenlik dönemi (middle-adolescent) örneklemleri üzerinde güvenirlilik ve geçerlik çalışmaları incelenmiştir.

Araştırmanın Amacı: Bu çalışmanın amacı, orta yaş ergenler (middle-aged adolescents) üzerinde ABABE'nin psikometrik özelliklerini belirlemektir. Bir diğer amaç ise, ergenler üzerindeki farklı etkileri bulunan anne ve babanın birlikte ele alınması yerine, yabancı yazın alanının da belirtildiği

gibi anne ve baba bağlanmasını ayırarak değerlendirmektedir. Böylelikle, ABABE'nin orta-ergenlik dönemindeki gençlerin, anne-baba ve akran ilişkilerinde güvenli bağlanma duygularını belirlemek amacıyla kullanılabilecek bir ölçme aracı olabileceğini göstermek çalışmanın genel amacıdır.

Araştırmanın Yöntemi: Çalışmaya, yaşları 14 ve 18 arasında değişen, Ankara'da orta öğrenimlerine Anadolu Lisesi ve Normal Lisede devam eden 315 öğrenci katılmıştır. ABABE'nin uyarılma çalışmasında kapsamlı ölçme yöntemleri üç farklı aşamada gerçekleştirilmiştir. Birinci aşamada, faktör yapıları ve iç tutarlılık katsayıları incelenmiştir. İkinci aşamada test-tekrar test ilişkisi ve üçüncü aşamada ise ölçüt bağımlı geçerliliği incelemek amacıyla ABABA'nin benlik saygısı ve iyi oluş ile ilişkisi test edilmiştir. Çalışmada kullanılan ölçme araçları, Anne-Baba ve Akran Bağlanması Envanteri, Benlik Saygısı Ölçeği ve Olumlu ve Olumsuz Duygu Ölçeği'dir. Açımlayıcı faktör analizi, temel bileşenler faktör analizi yöntemi ile incelenmiş ayrıca ölçme modelinin kabul edilebilir uyum iyiliği istatistikleri (goodness of fit statistics) üretip üretmediğini test etmek amacıyla doğrulayıcı (confirmatory) faktör analizi kullanılmıştır. Uyum iyiliği istatistikleri, herhangi bir modelin bir bütün olarak veri tarafından kabul edilebilir bir düzeyde desteklenip desteklenmediğine karar verme olanağı tanımaktadır. Yaygın olarak bilinen uyum istatistiği Kay-kare'dir (χ^2) ve bir modelin anlamlı olarak kabul edilebilmesi için χ^2 değerinin anlamlı çıkmaması beklenir. Ancak χ^2 değeri örneklem büyüklüğüne duyarlı olduğu için bir çok farklı uyum istatistiği üretilmiştir. Bunlar arasında en yaygın olarak kullanılanlar; Uyum İyiliği İndeksi (Goodness Of Fit Index, GFI), Düzeltilmiş Uyum İyiliği İndeksi (Adjusted Goodness Of Fit Index, AGFI), Karşılaştırmalı Uyum İyiliği İndeksi (Comparative Fit Index, CFI), Yaklaşık Hataların Ortalama Karekökü (Root Mean Square Error of Approximation, RMSEA) ve Standart Ortalama Hataların Karekökü (Standardized Root Mean Square Residual, SRMR) olduğu görülmektedir.

Araştırmanın Bulguları: Yapılan ilk doğrulayıcı faktör analizi sonuçları, ABABE'nin orijinal yapısının (Zero Model M_0) ürettiği uyum iyiliği değerlerinin kabul edilebilir değerlerin altında olduğunu göstermektedir. Analiz stratejisi gereği, ABABE'nin faktör yapılarını belirlemek amacıyla yapılan açımlayıcı faktör analizi sonuçları, anne, baba ve akran bağlanmasında üç faktörlü bir yapının olduğunu ortaya koymaktadır. Ancak her bir alt boyutta, annede 19, babada 24 ve akran'da 24 olmak üzere farklı madde sayıları gözlenmiştir. ABABE'nin üç alt boyutunu karşılaştırma olanağı sağlamak ve boyutlar arasında bir paralel form oluşturmak için anne, baba ve akran boyutunda, 45'in altında faktör yükü olan maddeler boyutlardan çıkarılmıştır. Sonuçta her üç boyutta (anne, baba ve akran), 18 maddeden oluşan kısa bir form elde edilmiştir. Ortaya çıkan bu yeni formun ölçme modelinin, uyum iyiliği değerleri ikinci doğrulayıcı faktör analizi ile test edilmiştir (Alternative Model, M_1). İkinci

doğrulayıcı faktör analizi sonuçları, uyum iyiliği değerlerinin kabul edilebilir düzeyde olduğunu göstermektedir. Bunun yanı sıra, ortaya çıkan yeni kısa form, bağımsız bir örneklem üzerinde test edilmiş, sonuçlar uyum iyiliği değerlerinin aynı şekilde kabul edilebilir düzeyde olduğunu ortaya koymuştur. İkinci aşamada gerçekleştirilen test-tekrar test sonuçları, anne ve baba toplam puanlar ilişkisinin yüksek olduğunu göstermektedir. Buna karşın akran bağlanması toplam puanları ve alt ölçekleri ile anne ve baba bağlanması alt ölçekleri test-tekrar test ilişkisinin orta düzeyde olduğu görülmektedir. Üçüncü ve son aşamada ise, ABABE'nin Benlik Saygısı ve Olumlu-Olumsuz Duygu Ölçeği ile ilişkisi ölçüt bağıntılı geçerlik olarak incelenmiştir. Sonuçlar, ABABE'nin Benlik saygısı ve olumlu duygu ile olumlu yönde, olumsuz duygu ile olumsuz yönde bir ilişki gösterdiğini ortaya koymaktadır.

Araştırmanın Sonuçları ve Öneriler: Genel olarak, çalışmanın ilk bulguları, ABABE'nin orijinal yapısının, orta yaş ergenler için kullanılabilir bir ölçme aracı olmadığını ortaya koymuştur. Ancak bir sonraki analizler sonucunda ortaya çıkan 18 maddelik yeni kısa formun bu yaş ergen grupları da anne, baba ve akran bağlanmasını değerlendirmeye yönelik bir ölçme aracı olduğunu göstermektedir. ABABE'nin yeni kısa formunun, orta öğrenimlerine devam eden orta yaş ergenlerin, anne-baba ve akranlarına bağlanma duygularını değerlendirmeye yönelik araştırmalarda kullanılabilir bir ölçme aracı olduğu görülmektedir. Bu çalışmanın sonuçlarının, Türkiye'de bağlanma kuramı ile sonraki süreçte gerçekleştirilecek çalışmalara ve geliştirilebilecek ölçme araçlarına kaynaklık edebileceği düşünülebilir.

Anahtar Sözcükler: Ergenlik, bağlanma, anne-baba, akran, geçerlilik

From Interpretations to Graphical Representations: A Case Study Investigation of Covariational Reasoning

Önder Köklü*

Elizabeth Jakubowski**

Suggested Citation:

Koklu, O., & Jakubowski, E. (2010). From interpretations to graphical representations: A case study investigation of covariational reasoning. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 40, 153-172.

Abstract

Problem Statement: The idea of change – both how things change and at what rate things change with respect to each other – is fundamental to a study of calculus, which is a critical course for students majoring in mathematics, sciences, engineering, business and several other majors. A lack of ability to reason about change in continuously changing functional relationships may cause difficulties in learning basic calculus concepts such as limits and derivatives. Despite a variety of research studies that emphasized the effects of students' understanding of rate of change on understanding of calculus concepts such as limits, derivatives, and integrals, there is little information about how college students reason about continuously changing functional relationships.

Purpose of Study: The aim of this study is to explore, describe and analyze college students' covariational reasoning abilities. More specifically, this study investigates and provides a "thick" description, or which explains a behavior and its context, of how a college student uses understanding and reasoning to interpret a functional situation and uses these interpretations to demonstrate the covariation of two variables in graphical representation.

Methods: Case study design and techniques were used in this study to provide a thick description of Karl's thinking and reasoning processes in

*Asst. Prof. Dr., Adiyaman University Faculty of Education, okoklu@adiyaman.edu.tr

** Prof. Dr. Florida State University College of Education, ejakubow@coe.fsu.edu.

order to comprehensively understand his covariational reasoning. Data were obtained from a detailed examination of student's thinking and reasoning processes through the task based in-depth clinical interviews. Data obtained from students' verbal expressions and graphical representations were analyzed in light of the theoretical lens. The "Covariation Framework" provided a skeletal structure for the description and interpretation of findings.

Findings and Results: Analysis of data disclosed that conceiving of a functional situation statically leads to difficulties in coordinating the continuously changing rate of change over the entire domain. Students' strong procedural tendency hinders meaningful interpretations and reasoning. Lack of transformational reasoning appears to prevent forming an image of the dynamically changing event and foster dependence on the procedural steps. Reasoning based on irrelevant details and arguments leads to either erroneous or pseudo-analytical conclusions about the simultaneous changes of two variables.

Conclusion and Recommendations: Instead of introducing the concept of function as correspondence, which is more traditional and concentrated around the application of certain rules and formulas, introducing the functions as covariation will be more helpful for students to develop a better conceptual understanding. Utilizing computer technology such as dynamic software in classroom instruction may provide more visual representations in order to enhance students' conceptualization of the changing nature of functions.

Keywords: Covariational reasoning, change, simultaneous changes, dynamic events.

Functions represent relationships between varying quantities, and, as Tall (1997) stated, "one purpose of function is to represent how things change" (p.1). Studies have revealed that students' underlying conceptions of function play an important role in coordinating the simultaneous changes of variables necessary to reasoning dynamically. Research studies (Carlson & Oehrtman, 2005; Monk, 1992; Confrey & Smith, 1995; Thompson, 1994b) agreed that dynamic conceptualizing of functions is essential to be able to coordinate simultaneous changes of variables. Monk (1992), for example, showed that students who had *point-wise* conceptions of functions, which is a static conception, had difficulties in describing patterns of change in the value of a function that results from a pattern of change in the values of the input variables. On the other hand, students who had an *across-time* conception, which is a dynamic conception, were able to describe the simultaneous changes of dependent and independent variables.

Moreover, Carlson and Oehrtman (2005) pointed out that a student with a *process view* of function can imagine the whole situation as it is happening at one time, rather than calculating every single pairs of input and output values. Thompson (1994b) has also supported this idea by stating that the *process view* of function allows one to imagine an expression rapidly evaluating itself.

Considering the studies of Breidenbach et al (1992), Dubinsky (1991) and Dubinsky & Harel (1992), Carlson and Oehrtman (2005), the process view of function is important and necessary to represent and interpret dynamic functional situations.

Covariational reasoning is defined by Carlson, Jacobs, Larsen, and Hsu (2002), as “cognitive activities involved in coordinating two varying quantities while attending to the ways in which they change in relation to each other” (p.354). Another definition comes from Saldanha & Thompson (1998). According to these researchers, it is “...holding in mind a sustained image of two quantities’ values (magnitudes) simultaneously” (p.298). It has been shown that these reasoning abilities are necessary for analyzing, interpreting and representing the patterns of change in continuously changing dynamic events. (Carlson et al, 2001; Monk, 1992).

According to Carlson & Oehrtman (2005) being able to imagine changes of variables simultaneously in a functional relationship is difficult, because it requires one to be able to picture the whole process at once and to be able to imagine running through several input and output pairs simultaneously, instead of performing specific computations for every input value. Thompson (1994b) expressed this ability as imagining the expression evaluate itself very rapidly over a continuum of the possible domain. This ability to imagine the whole process as it happens, all at once, has been linked to students’ underlying function conceptions.

Carlson (1998) introduced the covariation framework as a result of several investigations in which behaviors of undergraduate students were identified as they responded to tasks that involved interpreting and representing dynamic function situations. Their mental actions and corresponding behaviors were classified and defined by Carlson et al (2002) as shown in Table 1.

Thompson (1994a) mentioned the relationships between the accumulation of quantities and accruals that are constructed by an accumulation of quantities. According to Thompson, the development of mature images of rate involves the schematic coordination of these relationships. He then explained the stages of the development of images of rate as follows:

- Image of change in some quantity (e.g., displacement of position, increase in volume),
- A loosely coordinated images of two quantities (e.g., displacement of position and duration of displacement),
- An image of the covariation of two quantities so that their measures remain in constant ratio. (p.5)

According to Thompson, constructing an image of the covariation of two quantities is the last stage of the development of a mature image of rate. Moreover, Saldanha and Thompson (1998) investigated the images of covariation. They stated that images of covariation are also developmental and this development is defined in following stages:

- Coordination of two quantities’ values – think of one, then the other, then the first, then the second and so on.
- Construction of an operative image of covariation in which a person imagines both quantities having been tracked for some duration. (p.2)

Table 1*Mental Actions and Corresponding Behaviors*

Mental Action	Description of Mental Action	Behaviors
Mental Action (MA1)	An image of two variables changing simultaneously. (Coordinating the the value of one variable with changes in the other.)	<ul style="list-style-type: none"> Labeling the axes with verbal indications of coordinating the two variables (e.g., y changes with changes in x)
Mental Action (MA2)	A loosely coordinated image of how the variables are changing with respect to each other. (Coordinating the direction of change of one variable with changes in the other.)	<ul style="list-style-type: none"> Constructing an increasing straight line Verbalizing an awareness of the direction of change of the output while considering changes in the input.
Mental Action (MA3)	An image of amount of change of the output variable while considering changes in fixed amounts of the function's domain. (Coordinating the amount of change of one variable with changes in the other.)	<ul style="list-style-type: none"> Plotting points/ constructing secant lines. Verbalizing an awareness of the amount of change of the output while considering changes in the input.
Mental Action (MA4)	An image of rate / slope for contiguous intervals of the function's domain. (Coordinating the average rate of change of function with uniform increments of change in the input variable.)	<ul style="list-style-type: none"> Constructing contiguous secant lines for the domain. Verbalizing an awareness of the rate of change of the output (with respect to the input) while considering uniform increments in the input.
Mental Action (MA5)	An image of the continuously changing rate over the entire domain. (Coordinating the instantaneous rate of change of function with continuous changes in the independent variable for the entire domain of the function.)	<ul style="list-style-type: none"> Constructing a smooth curve with clear indications of concavity changes. Verbalizing an awareness of the instantaneous changes in the rate of change for the entire domain of the function (direction of

concavities and inflection points are correct)

Saldanha and Thompson (1998) also referred to the unpublished study of Coulombe and Brenson describing the development of the construction of covariation. Saldanha and Thompson (1998) stated that Coulombe and Brenson suggested the following properties to be involved in the concept of covariation.

- The identification of two data sets
- The coordination of two data patterns to form associations between increasing, decreasing, and constant patterns
- The linking of two patterns to establish specific connections between data values
- The generalization of the link to predict unknown data values.

As Yesildere and Türnüklü (2008) stated, the acquisition of certain mathematical skills is important. Although students' knowledge about the rate of change have been widely studied (Noble, Nemirovsky, Wright, & Tierney, 2001, Hauger, 1995; Hauger, 1997; Kaput, 1992; Monk & Nemirovsky, 1994; Orton, 1983) it is not quite clear how students represent and interpret simultaneous changes of variables in dynamic events. Believing that there is still a need to gain new information about how students coordinate the changes of two variables simultaneously in dynamic events, this study builds on existing studies in the literature by incorporating the notion of exploring, describing and analyzing a college student's covariational reasoning. The specific questions that guided the research are the following. (I) How do college students interpret simultaneous changes of two variables in functional situations? (II) How do college students use their interpretations to represent the covariation of two variables in a graphical context?

Method

The constructivist theory of learning has increasingly influenced mathematics education research. Since constructivist methodology is primarily concerned with a deeper understanding of an individual's reasoning, perspectives and purposes, case study design and techniques were used in this study to provide a thick description about Karl's thinking and reasoning processes in order to understand his covariational reasoning comprehensively. More specifically, through task-based clinical interviews, our purpose in this study is to elicit in-depth information about Karl's interpretations of the information in dynamic functional events and his use of this information to make graphical representations of simultaneous changes of variables.

A case study is defined by Stake (1995) as "the study of particularity and complexity of a single case, coming to understand its activity within important circumstances" (p. xi). Merriam (1998), on the other hand says "A qualitative case study is an intensive, holistic description and analysis of a single instance, phenomenon or a social unit" (p. 27).

Participant

Karl, who is majoring in mechanical engineering, was a high performing student in the "Calculus with Analytic Geometry III" class. This study took place in his second year in college. He stated that he was in the honors program in high school. He took Calculus-I and Calculus-II classes before taking the Calculus-III. He said that since he got good scores on both the Advanced Placement Test and SAT, he did not have to take College Algebra and Pre-Calculus courses. He volunteered among 7 high performing students whose average grades on tests and quizzes were "A." The course instructor provided us a list representing students' average grades. Course content consists of functions with multiple variables and their graphical representations, vectors, partial derivatives, optimization, multiple integration, coordinate systems (polar, spherical, and cylindrical), curves, line integrals, flux integrals divergence theorem.

Data Collection

Data were collected through in-depth clinical interviews with the selected student while he was working in a problem-solving situation. Interviews focused on exploring the student's thinking and the reasoning processes he used as he attempted to represent and interpret simultaneous changes of two variables in functional events. Since clinical interviews are the major sources in case studies, thinking and reasoning interviews provided most of the raw data in this study. Interviews were semi-structured and task-based. All interviews were audio taped and transcribed to record students' answers to the given questions. In order to improve the quality of the data representing the participant's thinking and reasoning, the participant was always urged to use a think aloud process. The participant was encouraged to verbalize everything while working on the tasks. Probing questions were used to encourage sufficient verbalization of the participant's thoughts.

Interview Tasks

Interview tasks consisted of a series of problems designed to explore students' thinking and reasoning about simultaneous changes of two variables when they attempt to represent and interpret dynamic functional situations. Tasks were adapted from the literature. Rather than plotting individual points in a given scaled coordinate system, the student was asked to produce qualitative graphs of a concrete situation and was asked to view them globally instead of point-wise, as explained by Monk (1992). In other words, the participant was asked to attend to the entire graph as an expression of the relationship between two simultaneously changing variables and express that relationship in words rather than numbers. Our rationale for probing the constructing of a qualitative graph of a given situation was to understand the participant's thought process, such as how he constructed an image and made transformations on that image when making a transition from a given

situation to graphical representation of that situation by coordinating two variables changing in tandem. Each task was presented on a separate piece of paper. Each task was first verbally presented to the subject, and then the subject was given time to read the task by himself. In representation tasks, students were given a dynamic functional situation and asked to represent this situation in a graphical context. In these tasks all situations were presented verbally and some of these with accompanying figures and shapes. "Representing the given situation in graphical context" meant constructing a graph of that situation. In other words, the participant was asked to attend to the entire graph as an expression of the relationship between two simultaneously changing variables and express that relationship in words rather than numbers.

Analysis

Considering our research focus, qualitative data analysis techniques were employed. Karl's verbal and graphical responses were examined in order to obtain valuable information about his thinking and reasoning processes. We used coding procedures after collecting the interview data in order to compare one unit of information with the next in looking for, as Merriam (1998) mentioned, "recurring regularities" (p.183). According to Bogdan and Biklen (1998), collected data involve these regularities in the form of repeated actions, words and phrases. Our goal in the analysis was to relate the observed regularities in data to existing literature by using prior categories which are compatible with the purpose and theoretical frame. Specifically, Carlson's (1998) classification of 'mental actions and corresponding behaviors' allowed us to analyze Karl's covariational reasoning.

Findings

Several tasks have been posed in order to discover the mental actions accompanying a collection of common behaviors exhibited while responding to these tasks. In these tasks, Karl was given functional situations and was asked to construct graphs representing the situations.

Circular Ripple Problem

Imagine a pebble is thrown into a lake, creating a circular ripple that travels outward at a constant speed. Sketch a graph that represents the area, A , of the circle as a function of time that has passed since the ball hit the lake.

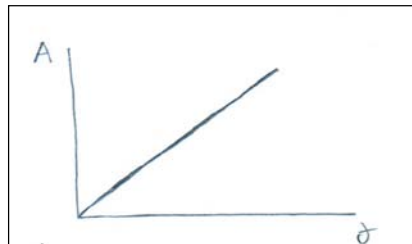


Figure 1. Task 1

Figure 2. Karl's first graph of

Area vs. Time

Karl: I am just thinking of a calm lake. I am throwing the pebble in to the lake and circle is getting larger and larger. It is increasing in every single direction and the whole circle is expanding...the circular ripple travels outward at a constant speed. So I can say that change in r (Δr) over change in time (Δt) would be constant, because in every moment it travels at a constant speed.

Interviewer: Could you explain why the area is increasing constantly?

Karl: As I said before the r [radius] is expanding constantly so the circle is growing at a constant rate...or it is not...let me think about it...the area is $A = \pi r^2$...when we plug in a different r every time, the increase in the area would not be constant.

Karl's first attempt indicated that the situation did not automatically evoke an image of how the area of a circle changes over fixed intervals of time. Even though he generated an initial image in his mind by stating "I am just thinking a calm lake. I am throwing the pebble in to the lake and the circle is getting larger and larger. It is increasing in every single direction and the whole circle is expanding...the circular ripple travels outward at a constant speed," he did not accurately manipulate this image through the use of mental transformations. Therefore he automatically presumed an inappropriate linear relationship between change in radius of the circle and change in the area of the circle. Considering the covariational framework (figure 1) his behaviors suggested that his mental actions are limited to coordinating the direction of change (MA2). After this immediate decision, he wanted to use the formula of the area of a circle to verify the relationship between radius and the area. Then he wanted to sketch a radius vs. area graph. Afterwards he plugged several integers in to " r " and plotted the corresponding area in the graph. After couple plotting, he just connected the plotted points and provided the following graph shown in figure 3. Then he stated:

Karl: You see the slopes of the lines [contiguous line segments]...they are different, and it is getting steeper and steeper...so the area increases more and more every time. But as a matter of fact the area depends on r^2 ...I mean it should be a parabola like in $y = x^2$ [giving an example]

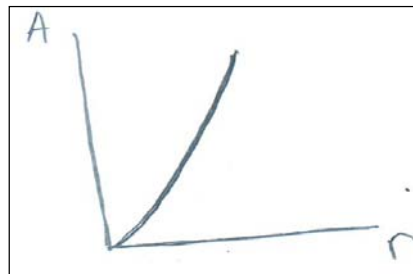
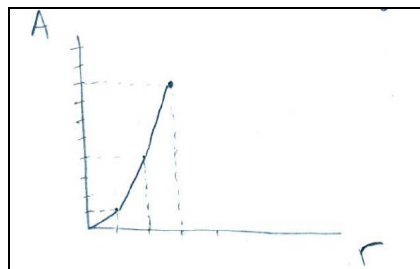


Figure 3. Karl's graph of area vs. radius

Figure 4. Karl's second graph

Then he changed his graph to a smooth curve as seen in the figure 4 by claiming: "...but as a matter of fact the area depends on r^2 ...I mean it should be a parabola like in $y = x^2$ ". Then he said that since the relationship between time and radius is linear, graphs of radius vs. area and time vs. area should be the same.

Although his later behaviors, such as constructing adjacent line segments with different slopes, may appear to be evidence of coordinating the amount of change in area with respect to change in radius, it is observed that his reasoning did not result from generating an image consisting of successive pictures in mind. Rather he followed a procedural method, by plugging in values to obtain correspondent values. Here we observed a lack of transformational reasoning. Simom (1994) explains the idea behind transformational reasoning as the ability of students to create a mental model that can be manipulated to visualize and understand relationships. A strong procedural tendency, such as looking for a formula to recognize the precise relationship, evidence of a lack of a transformational relationship. Although Karl's earlier statements, such as "I am just thinking of a calm lake. I am throwing the pebble into the lake and the circle is getting larger and larger. It is increasing in every single direction and the whole circle is expanding" provided evidence that he was forming an image of successive pictures of expanding circles, he obviously did not manipulate this image in order to make appropriate transformations resulting in the coordination of simultaneous changes of the variables of area and time; instead, he looked for a formula and used procedural methods.

It is also observed that Karl conceived this functional situation as static rather than dynamic. Plugging values into a formula to receive corresponding values, and plotting points and constructing contiguous line segments (figure 3), are some of the evidences of static conceptions. These behaviors also indicated that he was coordinating the amount of change in the area of the circle (MA3) (table 1). In the same way, Karl's statement "You see the slopes of the lines [contiguous line segments]...they are different, and getting steeper and steeper...so the area increases more and more every time" indicated his awareness of the rate of change of the output while considering the uniform increments of the input (MA4) (table 1).

Although he finally provided a smooth curve, this behavior did not demonstrate that his reasoning was based on understanding that the continuous changes in the area for the entire domain of the function resulted from smaller and smaller refinements of the average rate of change. His verbal statements such as "...but as a matter of fact the area depends on r^2 ...I mean it should be a parabola like in $y = x^2$ " indicated that he directly recognized the specific shape of a graph for the formula. These types of behaviors are classified as pseudo-analytical thoughts (Carlson et. al., 2002).

Balloon Problem

Imagine air is pumped into spherical balloon at a constant rate. Sketch a graph that represents the radius of the balloon as a function of the amount of air in the balloon.

Figure 5. Task-2

After reading the task Karl starts to think out loud. At the beginning he seemed little confused.

Karl: I have radius and...as air is being pumped in at a constant rate....since it is a spherical balloon, the radius is going to be changing at a constant rate as well...and I am just trying to think how it goes...but...I think the radius is going to increase faster at the beginning and slow down when it gets to the final volume. That it is how I visualize it.

Interviewer: Why do you think that the radius increases faster at the beginning and then slows down?

Karl: Because initially when you blow up the balloon, there is more space so the radius increases faster. Then, when it gets to the end, it [radius] is still increasing, but slowly, since there is not much room....

Interviewer: What about if there is no restriction of space? In other words, you can pump as much air into the balloon as you want. Do you think change in the radius would happen as you stated?

Karl: I am not hundred percent sure...let me think....I think I'm confusing myself [smiling]...I am going to write down the volume function [he wrote $V=4/3 \pi r^3$]. This is much clearer...since $4/3$ and π are constants, volume depends on r^3 .

After his statements he basically plugged several r values into the volume formula and came up with corresponding values of volume and sketched a graph of adjacent line segments by plotting the points and connecting these points.

First he made an effort to visualize situation in his head. He automatically presumed a linear relationship between the volume and the radius by claiming "...since it is spherical balloon, the radius is going to be changing at a constant rate as well..." This behavior suggested that his mental actions are limited to the coordination of direction of change (MA2) (table 1). Then he constructed an image of radius, increasing faster at the beginning and increasing more slowly at the end, by stating: "I am just trying to think how it goes...but... I think the radius is going to increase faster at the beginning and slow down when it gets to the final volume. That is how I visualize it." Construction of this image suggested that he was mentally estimating the rate of change in radius over equal intervals of the amount of air pumped into the balloon. But his following statements such as "...there is more space so the radius increases faster..." and "...it [radius] is still increasing, but slowly, since there is not much room...." suggested that his reasoning was not built upon mental actions of coordinating the rate of change of radius over small intervals of the amount of air pumped in. He was not observed conceptually coordinating the rate of change in radius; rather he was distracted by communicating irrelevant details in the image he generated, such as "...there is more space" and "...there is not much room...."

However, Karl was seen to be unsatisfied with his reasoning and conclusion. Then he wanted to apply a more analytical method to figure out how radius changes

with respect to the amount of air in the balloon by applying the volume formula. This behavior again suggested a lack of transformational reasoning. Although Karl's earlier statements "I am just trying to think how it goes...but... I think the radius is going to increase faster at the beginning and slow down when it gets to the final volume. That is how I visualize it" provided evidence that he was forming an image of successive pictures of an expanding radius as air pumped into a balloon. The mental transformations that he made to the object were only based upon irrelevant information or details. His procedural tendency, such as looking for a formula, carrying out appropriate computations and collecting evidence through procedural methods in order to recognize the precise relationship, are evidence of a lack of transformational reasoning.

He basically stated that the volume of the spherical balloon depends on r^3 , then plugged in several r values to obtain corresponding volumes. These behaviors also suggested that Karl conceived this functional situation as static rather than dynamic. Plugging values into a formula to get corresponding values, and plotting points and constructing contiguous line segments, are some evidence of static conceptions. These behaviors also indicated that he was coordinating the amount of change in the area of the circle (MA3) (table 1).

Filling Water Problem

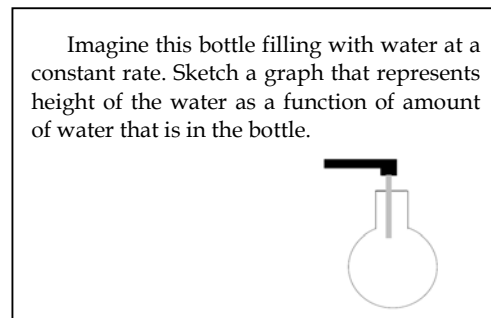


Figure 6. Task-3

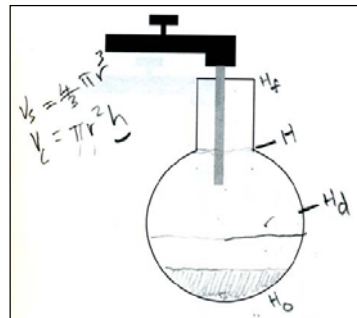


Figure 7. Karl's work

Upon reading the problem, he immediately searched for a formula while talking about the situation. First of all, he wrote down two formulas, one for the spherical part of the bottle ($V_s = 4/3 \pi r^3$) and the other for the cylindrical part ($V_c = \pi r^2 h$), as seen in figure 7. Then he claimed that volume of the spherical part of the bottle does not depend upon the height (h). Then he said that he could not utilize the formula to find out the relationship between amount of water and height, and stated:

Karl: The bottle actually is 3-dimensional, so we have specific volume. I'll do a graph of the amount of water and height. Instead of amount of water, I'll call it volume. Initially the volume will be at "0" for the height "0"...I will call it H naught or H_0 and ...here is the H final [placing marks on the figure for presenting the H_0

and H final]. Eventually we will have maximum volume. Since the radius of...ummm...at each different time there is a certain radius “ r ” and it is increasing. When the height goes up, volume slows down because the distance between the two walls is increasing as height goes up, and [this distance] eventually decreases because it is a sphere. So until it gets to the maximum distance between the two sides of the bottle, the rate of change in volume is going to be slowing down, and then it will be speeding back up. And then up here [showing the point where the cylindrical part starts] the rate of change would be constant. I mean, initially, the volume will increase at a slower and slower rate and then right here [showing the mid point of the sphere] if I call this H_d ...ummm...that would be the point where the graph changes from a greater rate of increase to...I mean it slows down and it speeds back up.

While he was expressing what he was thinking, he provided a graph (figure 8). Considering his verbal expressions such as “...as height goes up volume slows down” and the graph he provided (volume as a function of height), it is obvious that Karl constructed images of height changing constantly and volume changes at a varying rate. This image indicated that Karl has switched the roles of independent and dependent variables. After assuming the height is changing constantly he put the variable “Height” aside and introduced another variable and tried to explain the rate of change in volume by changes in the distance between the two walls of the bottle. However, his reasoning appeared to result in constructing an inappropriate graph.

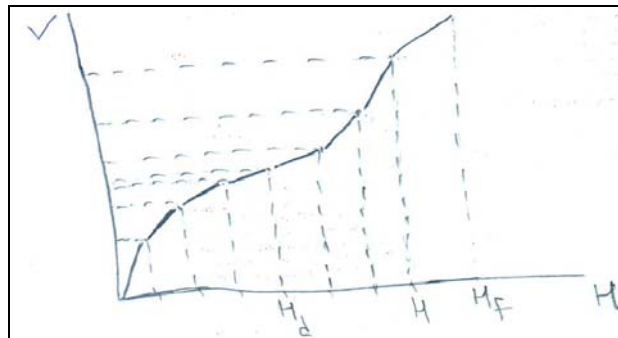


Figure 8. Karl's graph of volume vs. height

Interviewer: Could you provide us with more explanation about why the volume is slowing down?

Karl: As I said, the distance between walls is increasing so ...um...the rate of increase in volume is getting smaller and smaller until the midpoint of the bottle.

Interviewer: How do you know?

Karl: Since the distance is constantly increasing while the bottle is getting wider, the volume is increasing less and less.

Considering the statement “volume is increasing,” he appeared to be distracted by the image he constructed in his mind, and treated volume as height. In other words, his reasoning revealed that he was measuring the change of volume as changes in the level of water, which is in fact the measure of height. Although his behaviors such as constructing fixed intervals of height on the X axis and representing the relative changes in the amount of water or volume over these fixed intervals of height seemed to support the mental actions of coordinating the rate of change of the independent variable over the dependent variable, switching the roles of the variables of volume and height prevented him from providing the appropriate coordination of covariation between volume and the height. Overall, his behaviors such as his verbal expressions and the graph he provided suggested that inconsistencies in his reasoning prevented him from constructing accurate images or successive pictures of height changing at a varying rate as the amount of water in the bottle increases constantly. Therefore his mental actions were only limited to the coordination of the direction of change.

Conclusion and Discussion

Based on the research questions, the aim of this study was to explore, describe and analyze a college student’s covariational reasoning. Task-based interviews were used and data obtained from interviews were analyzed in light of Carlson et al (2002)’s “Mental Actions of Covariation” framework to identify the student’s thinking and reasoning processes as he attempted to coordinate simultaneous changes of variables in continuously changing dynamic functional situations. The participant exhibited various behaviors while applying his reasoning in order to analyze, interpret and represent dynamic functional events. In the light of the findings, the following assertions are made in this study.

Conceiving of a functional situation statically leads to difficulties in coordinating the continuously changing rate of change over entire domain. Karl’s thinking throughout the tasks demonstrates the evidence that functional situations are conceived as static rather than dynamic. Karl has been observed plotting points and obtaining adjacent linear lines for equal intervals in the domain rather than reasoning dynamically and disregarding the step by step evaluation of change. This static approach prevents him from evaluating the whole process dynamically as it is happening at once. It is clear that Karl initially focuses on actions when analyzing the relationship between variables. Analyzing the relationship on a point-to-point basis is an important consequence of the action view of functions. Karl’s graphic representations were frequently robust enough that he was easily plugging in values to obtain correspondent values. This view of function is static and did not allow him to coordinate changes in the dependent variable with the continuous changes in the independent variable.

Strong procedural tendency hinders meaningful interpretations and reasoning. In light of the analysis of the present case study, we noted that Karl showed his tendency to utilize algebraic expressions from time to time by either applying an algebraic expression, when available, or declared his need for an algebraic expression

to be able to construct a graph that represents the relationship between two variables. Karl was observed applying memorized procedures rather than conceptually coordinating the changes of two variables. This is an indication of a tendency towards using algebraic formulas, equations or memorized procedures to demonstrate a familiar situation in a graphical representation. For example, Karl used the formula of area of a circle to find out the respective amounts of area of a circle in task 1. He ended up an area vs. radius graph consisting of contiguous line segments (figure 3). Then he converted this graph to a concave up graph by claiming: "...but as a matter of fact the area depends on r^2 ...I mean it should be a parabola like in $y = x^2$." Although his construction of a smooth curve gave an appearance of engaging in coordinating the continuous changes in the dependent variable, his verbal expressions revealed that he actually did no more than apply a memorized rule. He was not able to provide a rationale for his construction. Carlson et al (2002) identified this kind of behavior as pseudo-analytical, as defined by Vinner (1997). According to Carlson, et al (2002) when students do not provide evidence that they possess an understanding that supports their behavior, this behavior is classified as pseudo-analytical behavior.

We noted that his inability or his reluctance to form an image of the dynamically changing events appeared to foster Karl's dependence on the procedural steps. For instance, in task 1 (figure 1), even though Karl first attempted to form an image of the situation, his verbal expressions indicated that the situation did not automatically evoke an image of how the area of a circle changes over fixed intervals of time. He presumed an inappropriate linear relationship between change in radius of the circle and change in the area of the circle. After this immediate decision, he wanted to use the formula of the area of a circle to verify the relationship between radius and the area. This was an indication of his inability to generate an accurate image of the situation. In response to task 2 (figure 5), Karl was also observed applying the volume of a sphere formula in order to figure out how radius changes with respect to amount of air in the balloon. He basically stated that the volume of the spherical balloon depends on r^3 , and then plugged in several r values to obtain corresponding volumes. Karl also attempted to use the volume of a sphere formula in task 3 (figure 6) as well. Upon reading the problem, he immediately searched for a formula. First of all, he wrote down two formulas, one for the spherical part of the bottle ($V_s = \frac{4}{3} \pi r^3$), and the other for the cylindrical part ($V_c = \pi r^2 h$), as seen in the figure 7. Then he said that he could not utilize the formula to find out the relationship, by claiming that volume of the spherical part of the bottle does not depend upon the height (h). This behavior also indicated his need for a formula and his dependence on procedural ways. Karl was observed applying memorized procedures rather than conceptually coordinating the changes of two variables. This is an indication of student's tendency towards using algebraic formulas or equation or memorized procedures to demonstrate a familiar situation through graphical representation.

We agree with Tall and Vinner (1981), who stated that students' restricted concept of function is rooted in the predominant use of functions given by algebraic formulas in traditional instructional methods. Leinhardt, Zaslavsky, and Stein

(1990) have also emphasized that the construction of graphs from algebraic formulas was highly dominant in traditional instructional methods.

Lack of transformational reasoning appears to prevent the formation of an image of the dynamically changing event and foster the student's dependence on procedural steps. Simon (1994) claims that the idea behind transformational reasoning is that students can create a mental model that can be manipulated to visualize and understand relationships. Carlson et al. (2002) have also emphasized in their study that mental actions involved in applying covariational reasoning are characteristic of transformational reasoning. In our investigation we observed examples of a lack of transformational reasoning.

For instance, in response to task 1 (figure 1), although Karl's statements such as "I am just thinking of a calm lake. I am throwing the pebble into the lake and circle is getting larger and larger. It is increasing in every single direction and the whole circle is expanding" provided evidence that he was forming an image of successive pictures of expanding circles; he obviously did not manipulate this image in order to make the appropriate transformations, resulting in the coordination of simultaneous changes of variables area and time. Instead, he looked for a formula that would fit into the situation and used procedural methods by plugging in values.

Similarly, in the second task (figure 5), Karl's procedural tendency to look for a formula, carry out appropriate computations and collect evidence through procedural methods in order to recognize the precise relationship, are evidence of a lack of transformational reasoning. This data supports a lack of transformational reasoning in the student's solution approach. Instead of manipulating the mental image through mental transformations in order to understand relationship, Karl looked for a geometric formula that would fit the situation.

Reasoning based on irrelevant details and arguments, leads to either erroneous or pseudo-analytical conclusions about simultaneous changes of two variables. Analysis of data also revealed several factors in the student's reasoning, which resulted in constructing inappropriate images of the relationship between two variables. Karl was observed communicating information that was not related to situations presented in the tasks. For example, while working on the balloon problem (figure 5), Karl claimed "... I think the radius is going to increase faster at the beginning and slow down when it gets to the final volume. That is how I visualize it. Because initially when you blow up the balloon, there is more space, so the radius increases faster than when it gets to the end. The [radius] still increases, but slowly since there is not much room." This expression suggested that his reasoning was not built upon mental actions of coordinating the rate of change of the radius over small intervals of the amount of air pumped in. He was not observed conceptually coordinating the rate of change in radius; rather, he was distracted by communicating irrelevant details in the image he generated, such as "...there is more space" and "...there is not much room." Similarly, in the ladder problem (figure 4.15) Karl claimed: "...When I think theoretically, the speed of top of the ladder should be constant, the same as the speed of the bottom of the ladder. Because when you move an object, speed should be the same at every part of that object," As seen in his verbal

expression, his reasoning is strictly bounded with irrelevant information when considering the situation in this specific task.

Implications

The present study does provide information about classroom instruction. However, we claim that the results of this qualitative case study provide valuable insights for instructional methods in classrooms, design and development of curricular activities and alternative assessment methods. An analysis of the data and findings revealed that even a high-performing calculus student has exhibited difficulties in constructing accurate images of a continuously changing rate of change in functions over domains. Taking this finding into consideration, this study suggests:

Since students' underlying function conception is said to be very important in the coordination of simultaneous changes of variables, the concept of function could be introduced to students in alternative ways in order to develop dynamic conceptions. More specifically, instead of introducing the concept of function as correspondence, which is a more traditional introduction and concentrated around the application of certain rules and formulas, it will be more helpful for students if the concept of function is introduced as covariation, which focuses on coordinating changes between input and output values, and emphasizes the changing nature of functions. This introduction would provide an outstanding foundation for students to develop a better conceptual understanding of functions.

Existing research also highlights the need for the creation of rich opportunities for students to understand relationships among variables. Utilizing computer technology such as dynamic software in classroom instruction may provide more visual representations in order to enhance students' conceptualizations of the changing nature of functions. This study also suggests the need for design and development of alternative curricular activities and alternative assessment tools for classroom teachers in order to develop students' covariational reasoning abilities and assess students' thinking and reasoning. In order to decrease the negative effects of overexposure of students to routine formula-dependent problems and tasks, providing students with alternative qualitative problems and tasks would be helpful.

References

- Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative Research for Education: An Introduction to Theory and Methods*. (3rd ed.). Boston: Allyn & Bacon.
- Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics*, 23, 247-285.
- Carlson, M. (1998). A cross-sectional investigation of the development of the function concept. In E. Dubinsky, A. H. Schoenfeld, & J. J. Kaput (Eds.), *Research in Collegiate Mathematics Education*, 1, (7), (pp. 115-162). Providence, RI: American Mathematical Society.

- Carlson, M., Jacobs, S., Coe, T., Larsen, S. and Hsu, E. (2002) "Applying Covariational Reasoning While Modeling Dynamic Events: A Framework and a Study". *Journal for Research in Mathematics Education*. Vol. 33, No. 5, 352-378.
- Carlson, M., Larsen, S., & Jacobs, S. (2001). An investigation of covariational reasoning and its role in learning the concepts of limit and accumulation. *Proceedings of the Twenty-Third Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Vol. 1, (pp.145-153), Snowbird, UT: PME-NA.
- Carlson, M., & Oehrtman, M. (2005). Key aspects of knowing and learning the concept of function. In A. Selden & J. Selden (Eds.), *MAA Online, Research Sampler*.
- Confrey, J., & Smith, E. (1995). Splitting, covariation, and their role in the development of exponential functions. *Journal for Research in Mathematics Education*, 26(1), 66-86
- Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 95-126). Boston: Kluwer Academic Publishers.
- Dubinsky, E., & Harel, G. (1992). The nature of the process conception of function. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy*, *MAA Notes*, 25, 85-106. Washington, DC: Mathematical Association of America.
- Hauger, G. S. (1995). *Rate of change knowledge in high school and college students*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco (April).
- Hauger, G. S. (1997). *Growth of knowledge of rate in four precalculus students*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago (March 24-28).
- Kaput, J. J. (1992). Patterns in students' formalization of quantitative patterns. In G. Harel & E. Dubinsky (Eds.), *The Concept of Function: Aspects of Epistemology and Pedagogy*, *MAA Notes*, Vol. 25 (pp. 290-318). Washington, DC: Mathematical Association of America.
- Leinhardt, G., Zaslavsky, O. & Stein, M. S. (1990). Functions, Graphs and Graphing: Tasks, Learning, and Teaching. *Review of Educational Research*, 1, 1-64.
- Merriam, S. B. (1998). *Qualitative research and case study: Applications in education*, Revised and expanded from case study research in education. Jossey-Bass Publishers, San Francisco.
- Monk, S. (1992). Students' understanding of a function given by a physical model. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy*, *MAA Notes*, 25 (pp. 175-194). Washington, DC: Mathematical Association of America.

- Monk, S., & Nemirovsky, R. (1994). The case of Dan: Student construction of a functional situation through visual attributes. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in Collegiate Mathematics Education. I. CBMS Issues in Mathematics Education* (pp. 139-168). Providence, RI: American Mathematical Society.
- Noble, T., Nemirovsky, R., Wright, T., & Tierney, C. (2001). Experiencing change: the mathematics of change in multiple environments. *Journal for Research in Mathematics Education*, 32(1), 85-108.
- Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics*, 14, 1-18.
- Saldanha, L., & Thompson, P. (1998). Re-thinking co-variation from a quantitative perspective: Simultaneous continuous variation. In S. B. Berensen & W. N. Coulombe (Eds.). *Proceedings of the Annual Meeting of the Psychology of Mathematics Education North America*. Raleigh, NC.
- Stake, R. (1995). *The art of case research*. Thousand Oaks, CA: Sage Publications.
- Tall, D. O. (1997) Functions and Calculus. In A. J. Bishop et al (Eds.), *International Handbook of Mathematics Education*, 289-325, Dordrecht: Kluwer.
- Thompson, P. W.(1994a). Images of rate and operational understanding of the fundamental theorem of calculus. *Educational Studies in Mathematics*, 26, 229-274.
- Thompson, P. W.(1994b). Students, functions, and the undergraduate curriculum. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in Collegiate Mathematics Education. I. CBMS Issues in Mathematics Education* (pp. 21-44). Providence, RI: American Mathematical Society.
- Yeşildere, S., & Türnüklü, E. (2008). An investigation of the components affecting knowledge construction processes of students with differing mathematical power. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 31, 151-169.

Bir Özel Durum Araştırması: Yorumlamadan Grafiksel Gösterime Kovaryasyonel Uslamlama

(Özet)

Problem Durumu: Değişim kavramı- matematiksel nesnelerin nasıl ve birbirine göre ne oranda değiştiği- matematik, fen, mühendislik, işletme vb. Alanlarda öğrenciler için önemli bir ders olan kalkülüs'ün temelidir. Sürekli değişim içeren fonksiyonel ilişkilerde, bu değişimler hakkında usamlama becerilerinin eksikliği öğrencilerin limit ve türev gibi temel kalkülüs kavramlarını öğrenmede zorluklar yaşamasına neden olmaktadır. Öğrencilerin "değişim oranı" kavramını anlamalarının limit, türev, integral

gibi diğer kalkülüs kavramlarını anlamları üzerindeki etkilerini araştıran bir çok çalışma yapılmış olmasına rağmen, öğrencilerin sürekli değişim içeren fonksiyonel ilişkilerde nasıl uslamlama yaptıkları konusunda yeteri kadar bilgi yoktur. Bu türden uslamlama bir ilişki içerisinde birbirine göre eşzamanlı olarak değişen iki değer koordinat edilmesindeki bilişsel aktiviteler olarak tanımlanır ve “kovaryasyonel uslamlama” olarak adlandırılır. Diğer bir tanımı ise iki değişkenin eşzamanlı değişimlerinin zihinde tutulan sürekli bir görüntüsüdür. Bu türden bir uslamlama sürekli değişim içeren dinamik olaylardaki değişim örüntülerini analiz etme, yorumlama ve göstermede zorunlu bir bilişsel aktivitedir. Bu bilişsel aktivite verilen fonksiyonel ilişkilerde her bir girdi ve karşılık gelen çıktıyı bulmak için yapılan hesaplamaların yerine bir çok girdi ve çıktı eşlerinin eşzamanlı olarak zihinde canlandırılması ve sürecin bir bütün olarak zihinde resmedilmesi yeteneğini içermektedir.

Araştırmanın Amacı: Bu çalışmanın amacı başarılı üniversite öğrencilerinin kovaryasyonel uslamlama yeteneklerinin bir özel durum çalışması ile tanımlanması, açıklanması ve analiz edilmesidir. Bir başka deyişle, bu çalışma başarılı bir üniversite öğrencisinin verilen bir fonksiyon durumunda değişkenlerin eşzamanlı değişimlerini yorumlarken kullandığı uslamlama yöntemleri ve bu yorumlardan yola çıkarak oluşturduğu grafik gösterimleri hakkında derin ve detaylı bir açıklama getirmeyi amaçlamaktadır.

Araştırmanın Yöntemi: Bu araştırma çalışmasında bir nitel araştırma yöntemi olan özel durum analiz teknikleri kullanılarak öğrencinin düşünme süreci ve bu süreç içerisindeki kovaryasyonel uslamlaması hakkında detaylı ve derinlikli bir açıklama getirilmiştir. Veriler problem tabanlı yüz yüze görüşmelerden elde edilmiştir. Yüz yüze görüşmeler öğrencinin verilen dinamik fonksiyon durumlarında var olan iki değişkenin eşzamanlı değişimlerini yorumlaması ve grafiksel gösterimlerini oluşturması sürecinde kullandığı uslamlamalarını detaylı olarak incelemek ve açıklamak için düzenlenmiş bir seri problemden oluşmaktadır. Öğrencinin sözlü ifadeleri ve grafiksel gösterimlerinden elde edilen veriler kavramsal mercek ışığında analiz edilmiştir. “Kovaryasyon çerçevesi” (Covariation Framework) bulguların açıklanması ve yorumlanmasında iskelet yapıyı oluşturmuştur.

Araştırmanın Bulguları: Verilerin analizi sonucunda, verilen bir fonksiyonel durumun statik olarak kavranması bir tanım kümesi boyunca sürekli değişen “değişim oranı”nın algılanmasında zorluklara neden olmaktadır. Katılımcının verilen fonksiyonel durumları değerlendirirken her bir girdi ve çıktı eşlemeleri için ayrı ayrı noktalar belirleyerek ve bu noktaları da birleştirerek grafiksel gösterimi yaptığı gözlenmiştir. Diğer bir deyişle öğrenci süreci veya fonksiyondaki ilişkiyi bir bütün olarak değil adım adım değerlendirmeler yaparak ortaya koymuştur. Bu da fonksiyonların dinamik değil statik olarak kavranmasından kaynaklanmakta ve fonksiyonel durum içerisindeki sürekli değişimin gösterilememesine neden olmaktadır. İşlemsel kurallara ve formüllere olan güçlü eğilim anlamlı

yorumlama ve uslamlamanın ortaya çıkmasını engellemektedir. Katılımcı zaman zaman iki değişken arasındaki ilişkinin grafiğinin çiziminde uygun cebirsel formülleri bulabildiği durumlarda kullanarak ve bulamadığı durumlarda da bir cebirsel formül veya ifadeye ihtiyaç olduğunu belirterek formüllere ve işlemsel kurallara olan güçlü eğilimini göstermiştir. *Dönüşümsel uslamlama* eksikliği öğrencinin dinamik olarak değişen bir durumun görüntüsünü zihinsel olarak oluşturmasını engellediği gibi kurallara ve formüllere bağımlılığını da artırmaktadır. Dönüşümsel uslamlama bir ilişkiyi anlamak ve zihinde canlandırmak için üzerinde dönüşümler yapılabilen bir zihinsel model yaratma yeteneğidir. Araştırmada öğrencinin bir çok durumda var olan ilişkiye ait zihinsel resimler veya görüntüler oluşturup ve bu resim ve görüntüleri zihinsel dönüşümler yaparak var olan ilişkiyi anlamak yerine geometrik formüller aradığı gözlenmiştir. Bulguların analizi sonucunda verilen durumla ilgili olmayan detaylar ve fikirler üzerine yapılandırılan uslamlamanın iki değişkenin eşzamanlı değişimleri hakkında hatalı veya ezbere dayalı sonuçlar çıkarılmasına neden olduğu görülmüştür.

Araştırmanın Sonuçları ve Önerileri: Fonksiyon kavramının geleneksel bir yöntem olan ve belirli kuralları ve formülleri uygulama üzerine kurulu "eşleme" veya "eşleştirme" yerine "eşzamanlı değişim" veya "kovaryasyon" olarak öğrenciye tanıtılmaya başlanması öğrencinin fonksiyon ve fonksiyonlar içindeki değişim kavramını daha etkin bir şekilde anlamasını sağlar. Bu yaklaşım ayrıca öğrencilerin fonksiyonları statik değil dinamik olarak kavramalarını sağlayacak ve "değişim oranı" "ani değişim oranı" ve "eşzamanlı değişim" gibi kavramları daha iyi anlamalarını sağlayacaktır. Bu çalışmada ayrıca öğrencilerin kovaryasyonel uslamlamalarının gelişimi için müfredatta alternatif aktivitelerin oluşturulması ve geliştirilmesi önerilmektedir. Buna paralel olarak öğrencilerin kovaryasyonel uslamlama becerilerinin öğretmen tarafından daha iyi anlaşılması ve belirlenmesi için alternatif ölçme ve değerlendirme yöntemlerinin de geliştirilip uygulanması önerilmektedir. Dinamik gösterimlere dayalı bilgisayar yazılımları veya paket programları öğrencilerin fonksiyon içindeki değişimi kavramsal olarak anlamaları için alternatif aktiviteler yaratmak ve daha görsel gösterimler sunmak için etkin bir şekilde kullanılabilir.

Anahtar Kelimeler: Kovaryasyonel uslamlama, değişim, eşzamanlı değişim, dinamik olay.

Well-Being of University Students Who Have Lived Abroad: The Case of Hacettepe University*

Hülya Şahin**

Yasemin Akman Karabeyoğlu***

Suggested Citation:

Şahin, H., & Karabeyoğlu Akman, Y. (2010). "Well-being of university students who have lived abroad: the case of hacettepe university." *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 40, 173-188.

Abstract

Problem Statement: In its broadest sense, well-being encompasses happiness, life satisfaction, fulfillment, stress, and the affective and cognitive evaluation of one's life. The concept of well-being includes humans' positive and negative emotional states, their life satisfactions, and domain satisfactions. It can be seen that variables which predict the well-being levels of many groups (young, adult, married, etc.) have been examined. Determining the variables that predict the well-being levels of university students who return to Turkey, will help to take measures which can make life easier for the aforesaid group.

Purpose of study: The purpose of this study is to reveal whether the well-being levels of the Turkish university students who have lived abroad for a while and continue their university education in Turkey vary with "Sex", "Time Spent Abroad", "Success at School", "Economic Conditions of the Family", "The Amount of Pocket-money", "Relations with Friends", "Relations with Instructors" and "Perceived Parental Attitude".

Methods: At Hacettepe University, in the 2007-2008 academic year, a Subjective Well-being Scale was applied on 111 students at the departments of German, English, and French Language Teaching, where the number of students who have lived abroad is the highest. Students who have lived abroad for a period of at least 1-3 years were accepted as

* The article was presented in 18 th Education Science Congress in Kuşadası-İzmir (1-3 October 2009).

** Asst. Prof. Dr., Mehmet Akif Ersoy University Faculty of Education, hsahin@mehmetakif.edu.tr

*** Prof. Dr. Mehmet Akif Ersoy University Faculty of Education, akman@mehmetakif.edu.tr

“students with experience abroad”. The data acquired were tested by t test and one-way analysis of variance for unrelated measurements.

Findings and Results: The findings of the study show that the well-being levels of university students who have lived abroad do not vary with sex, time spent abroad, perception of success at higher education, and the economic conditions of the families; while they do vary by parental attitudes, relations with instructors and friends, and pocket-money variables.

Recommendations: It is recommended for university students who have lived abroad and continue their university education in Turkey that services shall be provided to improve their interpersonal communication skills, their pocket money shall be regulated in order to meet their daily needs, and parents shall be provided with training on a democratic and tolerant attitude no matter how old their child is.

Keywords: Well-being, university student, living abroad, parental attitude, interpersonal relations

As of the moment individuals start to exist, they have to adapt to the environment in which they live. Efforts for adaptation cause some difficulties. Therefore, efforts aiming at understanding humans and improving adaptation focus rather on the problems experienced. In addition, it is possible to say that studies carried out in the psychological field from the 1990s do not solely aim at understanding negative experiences or eliminating their effects. In the past decade, such concepts as Happiness, Wellness, and Well-being have been research subjects for psychology.

In the literature it could be seen that the concept of well-being is described to include humans' positive and negative emotional states, life satisfactions, and domain satisfactions. In its broadest sense, well-being encompasses happiness, life satisfaction, fulfillment, stress and the affective and cognitive evaluation of one's life (Prieto, Diener, E., Tamir, Scollon & Diener, M., 2005). Today, researchers who study SWB, conceptualize it as a multi-faceted field which includes cognitive and affective sub-components. The affective components of SWB include negative and positive emotions. The affective component represents the ongoing part of SWB. The cognitive components of SWB, however, include evaluations regarding judgements of life satisfaction. The judgements generally represent life satisfaction or domain satisfaction. Domain satisfaction is usually centered on specific fields of life such as work, marriage, and leisure time. Life satisfaction in general includes cognitively-focused evaluations regarding life as a whole (Pavot & Diener, 2004).

SWB introduces a variety of means in evaluating the life quality of people. Within the emotional field, individuals may react with their experience, many positive emotions and few negative or unsatisfactory experiences. In the mental field, they may make judgements on life satisfaction and fulfilment in their lives because

everything is going alright in a person's life as long as s/he is in a subjective well-being state (Diener, Kesebir & Lucas, 2008).

Studies that have contributed to the concept of well-being have constituted a comprehensive definition of psychological adaptation. Ryff and Singer (1996) explain "well-being" as the state free of major psychological problems and suggest a theoretical model consisting of six components regarding the concept of well-being. These components are self-acceptance, establishing positive relationships with others, autonomy, controlling the environment, aim in life, and personal development (Aysa, McCarrey & Ark, 2001). This multi-dimensional model has integrated with positive functions, Erikson's stages of psychosocial development and Maslow's concept of self-realization. This model was put forward as a reaction against lack of theoretical fundamentals for the concept of well-being. Although six structures having intermediate relations with each other were defined by Ryff (1995), these were seen as adequately independent and separate structures (Aysa, McCarrey & Ark, 2001).

When the literature is examined, one can encounter many studies on differences between groups and the relations between well-being and sex (Shmotkin, 1990; Fujita, Diener, & Sandvik, 1991; Diener, Suh, Lucas, & Smith, 1999; Nolen-Hoeksema & Rusting, 1999; Lucas & Gohm, 2000; Holmes, 2001; Dost, 2004), age, education (Keyes, Shmotkin and Ryff, 2002), culture (Elliot, Chirkov, Kim & Sheldon, 2001; Quraishi & Evangeli, 2007), life satisfaction (Boschen, 1996, Greenspoon & Saklofske, 1998), irrational beliefs (Çivitçi, 2009), depression, anxiety, stress (Katja, Päivi, Terttu & Pekka 2002; Kazarian, 2005), and economic income (Diener and Biswas-Diener, 2002)

When the findings of the study are examined, it can be seen that well-being level does not vary with sex and income levels (Shmotkin, 1990; Fujita, Diener, & Sandvik, 1991; Diener, Suh, Lucas, & Smith, 1999; Nolen-Hoeksema & Rusting, 1999; Lucas & Gohm, 2000; Holmes, 2001; Dost, 2004; Diener and Biswas-Diener, 2002). On the other hand, it is observed that well-being level creates significant differences in life satisfaction, irrational beliefs, and competence in social relations (Boschen, 1996, Greenspoon & Saklofske, 1998; Diener & Seligman, 2002; Çivitçi, 2009). Keyes, Shmotkin and Ryff (2002), observed that optimal well-being increased as age, education, extraversion, and conscientiousness increased and as neuroticism decreased. While there are studies stating that well-being levels vary with cultural features (Quraishi and Evangeli, 2007), there are also others that claim the opposite (Elliot, Chirkov, Kim & Sheldon, 2001). Findings of the studies that examine the relation between the state of health or actions aiming at health protection and well-being (Katja, Päivi, Terttu & Pekka 2002; Kazarian, 2005), show that there is a high positive relationship between well-being and the behaviours in question. Gençöz and Özlale (2004) investigated direct and indirect effects of social support on psychological well-being. And they observed aid-related social support and psychological well-being association was partially mediated by experiencing fewer life stresses. Otherwise appreciation-related social support had a direct effect on psychological well-being.

In the literature, no study exists on the well-being levels of students who were born in another land or culture, lived there for a certain period and came to Turkey

to continue their education. However, problems still exist for the third generation of families who immigrated to Central Europe, although these young people seem to have adapted to the environment they live in (II. International Congress of European Turks within the context of Education and Culture, final declaration, 2009). According to the Specialization Commission Report of the eighth five-year development plan, "inequalities in education" is one of the important problems. In European countries such as Belgium and Germany, where the number of Turkish inhabitants is relatively high, it is noticeable that when compared to these countries' own young population, the ratio of Turkish students who can continue their higher education is very low.

The author studying "The Profiles of Young European Turks who Study at Higher Education Institutions in Turkey" (2008) established that, as compared with their peers who were born and grew up in Turkey, these young students spend more time studying at university than they do in secondary education and participate in fewer social activities. In the aforementioned study, the students expressed that they were content with their higher education in Turkey, and that they would be less successful if they continued their studies abroad.

In this study, the aim is to reveal whether the well-being levels of Turkish University students who have lived abroad for a while and continue their higher education in Turkey vary with "sex", "time spent abroad", "success at school", "economic conditions of the family", "the amount of pocket-money", "relations with friends", "relations with instructors" and "perceived parental attitude".

Method

Participants

In the 2007-2008 school year, bachelor's level students at Hacettepe University, Departments of German, English, and French Language Teaching, where the number of students who had lived abroad is the highest, were included in the study. Students who had lived abroad for at least a period of 1-3 years were accepted as "students with experience abroad". The research group consists of 111 students, 35 (31,5 %) of which are males and 76 (68,5%) females. The distribution of the research group students by departments and time spent abroad is given in Table 1.

Table 1*The Distribution of the Research Group by Departments and Time Spent Abroad*

Time		Department			Total
		German Language Teaching	English Language Teaching	French Language Teaching	
1-3 years	n	11	-	-	11
	%				9,9
4-6 years	n	14	-	-	14
	%				12,6
7-10 years	n	12	2	1	15
	%				13,5
11-13 years	n	22	-	-	22
	%				19,8
14 and more	n	47	1	1	49
	%				44,1
Total	n	106	3	2	111
	%	95,5	2,7	1,8	100

Table 1 shows that almost 96% of the students with experience abroad study at the Department of German Language Teaching. It can be seen that nearly 64% of the students have lived abroad for 10 years and more.

Instruments

Research data concerning the dependent variable were gathered by the "Subjective Well-being Scale" prepared by Tuzgöl Dost (2004). Data regarding independent variables were acquired with the personal information forms prepared by the researchers. The "Subjective Well-being Scale" is composed of 46 items, and includes personal judgements regarding life domains and positive and negative emotional expressions. The scale is in the form of a 5-point Likert scale using the terms "completely suitable", "mostly suitable", "partially suitable", "slightly suitable" and "not suitable at all". Twenty-six items of the scale are positive while 20 of them are negative. Negative expressions include the 2, 4, 6, 10, 13, 15, 17, 19, 21, 24, 26, 28, 30, 32, 35, 37, 38, 40, 43 and 45 items. The negative expressions are scored by inversion. The lowest score that can be reached on the scale is 46, and the highest 230. A high score indicates a high level of well-being.

Validity of the Subjective Well-Being Scale. As for the validity of the Subjective Well-Being Scale, for peer opinion, structural validity, item validity, and item-total correlations, the discriminant function of the test was checked by means of bottom-up 27% score groups as compared by t test. Factor loadings obtained at the end of the factor analysis of the scale items vary between 30 and 664. The bottom-up 27% score means obtained by applying the 46-item scale on 209 people in order to check the discriminant function of the test were compared by t test and the result was found meaningful ($t:26.88, p<.01$). Moreover, Beck Depression Inventory was used for parallel form validity. An opposite relation coefficient at 01 level and a value of -70 was found between the scores obtained from each scale. As a result, the scale has been concluded to be valid based on all validity methods estimated (Dost, 2004).

Reliability of the Subjective Well-Being Scale. The Cronbach Alfa reliability coefficient was found to be 93. While the test/re-test reliability coefficient obtained by applying the scale on 39 people and applying it again after two weeks was found to be 86 (Dost, 2004).

Data Analysis

The data acquired from the well-being scale were tested as to whether they meet normal distribution and homogeneity of variances which are among parametric statistical assumptions and were found to meet parametric statistical assumptions. Therefore, t test for independent sample groups was used in order to test the difference between two group means, and One-Way analysis of variance (ANOVA) to test the difference between means of more than two groups. In cases where variances are homogenous, LSD and Tukey multiple comparison tests were used in order to determine the cause of the difference between groups. Coefficient of error was taken as .05.

Findings and Results

Whether the "Well-Being Scale" score means of students with experience abroad vary with sex was analysed by t test. Accordingly, t test results of well-being scale score means of university students by sex are given in Table 2.

Table 2

t Test Results

Sex	n	\bar{X}	s	df	t	Sig.
male	35	175,7	22,5	109	-,341	,734
female	76	177,4	25,3			

When Table 2 is analysed, it can be seen that there is no significant difference between the students' well-being score means by variable of sex [$t(109)= -,341, p>.05$]. No significant difference by sex was observed in the university students' well-being levels.

Whether university students' well-being scores vary with time spent abroad variable was tested by ANOVA and the results are presented in Table 3.

Table 3

ANOVA Results

Source	SS	df	MS	F	Sig.
Between Groups	2247,114	4	561,778	,944	,442
Within Groups	58331,275	98	595,217		
Total	60578,388	102			

The results of analysis show no significant difference between the students' well-being scores by *time spent abroad* [$F(4-98)=,944$, $p>.05$]. In other words, well-being levels of the students do not vary with time spent abroad.

Whether university students' well-being scores vary with the perception of success at higher education was tested by ANOVA and the results are presented in Table 4.

Table 4

ANOVA Results

Source	SS	df	MS	F	Sig.
Between Groups	4493,091	3	1497,697	2,630	,054
Within Groups	60924,009	107	569,383		
Total	65417,099	110			

According to Table 4, there is no significant difference between the students' well-being scores by *the perception of success at higher education* [$F(3-107)=2,630$ $p>.05$]. Accordingly, it is observed that well-being levels of the students do not vary with the perception of success at higher education.

Whether university students' well-being scores vary with the economic conditions of the family was tested by ANOVA and the results of analysis are presented in Table 5.

Table 5

ANOVA Results

Source	SS	df	MS	F	Sig.
Between Groups	4927,807	5	985,561	1,702	,141
Within Groups	60233,866	104	579,172		
Total	65161,673	109			

The results of the analysis show no significant difference between the students' well-being scores by *the economic conditions of the family* [$F(5-109)=1,702$ $p>.05$]. In other words, well-being levels of the students do not vary with their perception regarding families' economic conditions.

Whether university students' well-being scores vary with the amount of *their pocket money* was tested by ANOVA and the results of analysis are presented in Table 6.

Table 6

ANOVA Results

Source	SS	df	MS	F	Sig.	Significant Difference
Between Groups	4617,665	2	2308,833	4,080*	,020	A-C
Within Groups	60544,007	107	565,832			
Total	65161,673	109				

* $P<.05$ (A=Adequate amount of pocket money, C=Inadequate amount of pocket money)

The results of the analysis show a significant difference between the students' well-being scores by *the amount of their pocket money* [$F(2-107)=4,080$ $p<.05$]. According to the result of the Tukey test applied in order to find the cause of this difference, the well-being levels of the students who find the amount "adequate" ($\bar{x}=183$) is higher than those who find it "inadequate" ($\bar{x}=168$). Accordingly, there is a difference in favour of students who find the amount adequate.

Whether university students' well-being scores vary with their relations with friends was tested by ANOVA and the results of analysis are presented in Table 7.

Table 7
ANOVA Results

Source	SS	df	MS	F	Sig.	Significant Difference
Between Groups	7901,131	2	3950,566	7,382*	,001	A-B
Within Groups	57260,541	107	535,145			
Total	65161,673	109				

*P<.01 (A=Quite good, B= Good)

According to Table 7, there is a significant difference between the students' well-being scores by *their friend relationships* [F(2-107)=7,382 p<.01)]. According to the result of the Tukey test applied in order to find the cause of this difference, the well-being levels of the students who express that their friend relationships are *quite good* (\bar{x} =183) is higher than those who say they are *good* (\bar{x} =168). The research finding shows that there is a difference in favour of students who describe their relationships as quite good.

Whether university students' well-being scores vary with their relations with instructors was tested by ANOVA and the results of analysis are presented in Table 8.

Table 8
ANOVA Results

Source	SS	df	MS	F	Sig.	Significant Difference
Between Groups	6242,684	2	3121,342	5,647*	,005	A-C, B-C
Within Groups	58590,032	106	552,736			
Total	64832,716	108				

*P<.01 (A=Quite good, B=Partially good, C= Not good at all).

The results of the analysis show that there is a significant difference between the students' well-being scores by *their relations with instructors* [F(2-106)=5,647 p<.01)]. According to the result of the LSD test applied in order to find the cause of this difference, the well-being levels of the students who express that their relations with instructors are *quite good* (\bar{x} =179) and *partially good* (\bar{x} =182) is higher than those who say they are *not good at all* (\bar{x} =165). The difference found is in favour of groups who describe their relations as quite good and partially good.

Whether university students' well-being scores vary with perceived parental attitudes was tested by ANOVA and the results of analysis are presented in Table 9.

Tablo 9

ANOVA Results

Source	SS	Df	MS	F	Sig.	Significant Difference
Between Groups	9165,019	4	2291,255	4,294*	,003	A-D
Within Groups	54959,898	103	533,591			
Total	64124,917	107				

*P<.01 (A= Democratic and Tolerant, D= Indifferent)

According to Table 9, there is a significant difference between the students' well-being scores by *the parental attitudes they perceive* [$F(4-103) = 4,294$ $p < .01$]. According to the result of the Tukey test applied in order to find the cause of this difference, the well-being levels of the students who describe their parents' attitudes as *democratic and tolerant, good* ($\bar{x} = 187$) is higher than those who say they are *indifferent* ($\bar{x} = 147$). The difference found is in favour of the group who describes their parents' attitudes as democratic and tolerant.

Discussion and Suggestions

The first of the problems dealt with in the study was whether the well-being levels of the students who have lived abroad for a while vary with their sex. In the analyses of the relevant literature, it has been found that sex is not a predicted variable for the well-being levels (Shmotkin, 1990; Fujita, Diener, & Sandvik, 1991; Diener, Suh, Lucas, & Smith, 1999; Nolen-Hoeksema & Rusting, 1999; Lucas & Gohm, 2000; Holmes, 2001; Dost, 2004), while it has been considered that especially female students may feel more under pressure in Turkey in comparison with their life abroad and this may adversely affect their well-being levels. Findings obtained confirmed this opinion as they were parallel with the literature. Taking into consideration the fact that students in the research group are university students, with research carried out on students of different age groups, the effect of sex on the well-being levels of students who have lived abroad can be generalized.

Research findings show no significant difference between students' well-being levels and the time spent abroad. Time spent abroad was emphasized in terms of the adaptation process of students who return to their country after an experience in another land or culture. It was thought that the longer the time spent abroad, the harder it may be to adapt to Turkey and that this situation may create a significant difference between well-being levels of the students. However, in the focus group discussions held with research group students, it was discovered that most of the students had returned to Turkey at least a year ago. Therefore, it could be asserted that students might have adapted to the environment in this period.

According to the reasearch findings, there is no significant difference in students' well being levels with regard to their success perception. This finding is in conflict with Dost's finding (2004) with which he established a difference in favour of

students who consider themselves successful. The author (2008) in a similar study, stated that students who return to Turkey for higher education spend more time studying than they do in secondary education. Based on this, it can be thought that students may attribute their failure to reasons other than academic inadequacy and hence the variable of academic success may not have created a difference in students' well-being levels.

No significant difference in well-being levels of the students who have lived abroad was found by their perception of the family's economic income, whereas it was found that well-being levels vary with their perception of pocket money. It was observed that the well-being levels of the students who find the amount of daily pocket money "adequate" is higher than those who find it inadequate. Research findings on income and well-being relations vary. Diener and Biswas-Diener (2002) observed that although the well-being levels of individuals with good income states are slightly higher than those whose income is not good, this difference does not constitute a meaningful indicator; in fact individuals who prefer material goals to other values as the aim in life are less happy. Nevertheless, Boschen (1996) and Shek (2003) have stated that income level predicts life satisfaction that is an indicator of well-being. As for this study, it could be asserted that a similar finding to Boschen (1996) and Shek's (2003) research was obtained taking into consideration the fact that students' well-being levels vary with their pocket money, but not with the economic conditions of the family. Well-being levels of the university students, who can be considered as young adults, vary with their own incomes and not with the family's.

When the well-being levels of the students are examined by their relationships with friends and instructors, it can be seen that the well-being states of the students who claim to have "quite good" relations with instructors are higher than those who claim their relations are "not good at all", and students who describe their relations as "partially good" have higher well-being levels than those whose relations are "not good at all". Similarly, it was found that students who describe their friend relationships as "quite good" have higher levels of well-being than those who describe them as "good". Diener and Seligman (2002), having dealt with well-being from the point of social relations, expressed that good social relations are necessary for well-being, but not adequate. In addition, Regina and Loke (2005), in their study, have found that well-being levels of female students whose interpersonal relations are good, are higher. It could be asserted that results obtained in the research appear to be parallel with the findings of the aforementioned researchers. When well-being levels are analysed in terms of parental attitudes, well-being levels of the students who describe their parents' attitudes as "democratic and tolerant" were found to be higher than those of the students who describe them as "indifferent". This finding is supportive to Dost's (2004) research findings. As emphasized in Ryff and Singer's theoretical model (1995), which they carried out regarding the well-being concept, and also accepted in this research, well-being levels of the students who can establish positive relations with others seem higher.

In order to improve interpersonal relations of students who have lived abroad and continue their higher education in Turkey, some services should be provided. Thus, a hand can be given to improve their relations with instructors and friends. Scholarship opportunities should be improved in order to arrange their pocket

money so that it could meet their needs. Parents should be provided with training on democratic and tolerant attitudes towards their children no matter how old they are. Researchers who will work with similar groups are advised to extend the time anticipated in order to accept that the students have lived abroad and to include the students within the scope of the study in the very first year they come to Turkey.

References

- Aysa, N. & Mccarrey M. (2001). " The Relation Between Well-Being Impostor Feeling and Gender Role Orientation Among Cnadian University Students", *The Journal of Social Psychology*, 141 (2), 218-232.
- Boschen, K.A. (1996). Correlates of life satisfaction, residential satisfaction and locus among adults with spinal cord injures. *Rehabilitation Counseling Bulletin*. 39(4),Jun.
- Ciarroch, J. & Scott, G. (2006). The link between emotional competence and well-being: a longitudinal study. *British Journal of Guidance & Counseling*, 34 (2), May.
- Çivitçi, A. (2009). Relationship between irrational beliefs and life satisfaction in early adolescents. *Eurasian Journal of Educational Research*, 9(37),91-109.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: three decades of progress. *Psychological Bulletin*, 125, 276-302.
- Diener, E. & R. Biswas-Diener (2002). Will money increase subjective wellbeing: a literature review and guide to needed research. *Social Indicators Research*, 57, 119-169.
- Diener,E., Kesebir, P. & Lucas, R. (2008). Benefits of accounts of well-being-for societies and for Psychological science. *Applied Psychology: An International Review*, 57, 37-53.
- Dost T.M. (2004).Üniversite öğrencilerinin öznel iyi oluş düzeyleri. [Well-being levels of university student]. Unpublished doctoral dissertation,, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Elliot, A.J., Chirkov, V.I., Kim, Y. & Sheldon K.M. (2001). A cross-cultural analysis of avoidance (relative to approach) personal goals. *Psychological Science*, 12, 505-510.
- Ferring D. & Boll T. (2009).Regulation of Subjective Well-being in Older Adults. Research Unit INSIDE, University of Luxembourg.
- FÖTED Almanya Türk Veli Dernekleri Federasyonu (2004). Retrieved january 2008 from <http://tuerkische-elternfoederation.de/0-Basina/seite6.html> .
- Fujita, F, Diener, E., & Sandvik, E. (1991). Gender differences in negative affect and well-being: The case for emotional intensity. *Journal of Personality and Social Psychology*,61,421-34.
- Gençöz T. & Özlale, Y. (2004). Direct and indirect effects of social support on psychological well-being. *Social Behavior and Personality*, 32 (5),449-458.
- Greenspoon, P.J & Saklofske, D.H. (1998). Confirmatory factor analysis of the multidimensional students'life satisfaction scale. *Personality and Individual Differences*, 25, 965-971.
- Holmes, D. S. (2001). *Abnormal psychology* (4. ed.), Boston, MA: Allyn and Bacon.

- Katja, R., Päivi, K.A., Terttu, M.T. & Pekka L. (2002). Relationships among adolescent subjective well-being, health behavior and school satisfaction. *Journal of School Health*. 6 (72), 243-249.
- Kazarian S.S. (2005). Family functioning, cultural orientation, and psychological well-being among university students in Lebanon. *The Journal of Social Psychology*, 145 (2), 141-152.
- Keyes, C. L. M., Shmotkin, D. & Ryff, C. D. (2002). Optimizing well-being: The empirical encounter of two traditions. *Journal of Personality and Social Psychology*. 82 (6), 1007-1022.
- Lucas, R. E. & Gohm, C. L. (2000). Age and sex differences in subjective well-being across cultures. In E. Diener & E. M. Suh (Eds.). *Culture and Subjective Well-Being*, 291-318.
- Manco, A. A. (2006). Forty years of Turks in Belgium (1966-2004). Problems, developments, changes. Institut de Recherche, Formation et d'Action sur les Migrations (I.R.F.A.M.), Retrieved December 2007 from <http://www.flwi.ugent.be/cie/IRFAM/amanco8.htm> .
- Nolen-Hoeksema, S. & Rusting, C. L. (1999). Gender differences in well-being. In D. Kahneman, E. Diener, and N. Schwarz (eds.). *Well-being: The Foundations of Hedonic Psychology*, 330-352, New York: Russell Sage Foundation.
- Pavot, W. & Diener, E. (2004). The subjective evaluation of well-being in adulthood: findings and implications. *Ageing International*, Spring, Vol. 29, No. 2, 113-135.
- Prieto, C.K., Diener, E., Tamir, M., Scollon, C. & Diener, M. (2005). Integrating the diverse definitions of happiness: a time-sequential framework of subjective well-being. *Journal of Happiness Studies*, 6, 261-300.
- Regina L. T. L. & A. J. T. Yuen Loke (2005). Health-promoting behaviors and psychosocial well-being of university students in HongKong. *Public Health Nursing*, May, Vol. 22 No. 3, 209-220.
- Shahe S. K. (2005). Family functioning, cultural orientation, and psychological well-being among university students in Lebanon. *Lebanon The Journal of Social Psychology*, 145(2), 141-152.
- Shek, D. T.L. (2003). Economic stress, psychological well-being and problem behavior in Chinese Adolescents with Economic disadvantage. *Journal of Youth and Adolescence*. 32, 4.
- Shmotkin, D. (1990). Subjective well-being as a function of age and gender: a multivariate look for differentiated trend. *Social Indicators Research*, 23, 201-230.
- OSYS (Preference Guide) (2007). Retrieved December 2007 from [ftp://dokuman.osym.gov.tr/\\$2007 OSYS-TERCIH-KILAVUZU/](ftp://dokuman.osym.gov.tr/$2007%20OSYS-TERCIH-KILAVUZU/) .
- Quraishi, S & Evangeli, M. (2007). An investigation of psychological well-being and cultural identity in British Asian female university students. *Mental Health, Religion & Culture*, May, 10(3), 293-300.
- T.C. Başbakanlık Devlet Planlama Teşkilatı (2001). Labor Market (Turkish People Living Abroad) Specialization Commission Report. Kitap No. 3006. Ankara: Author.
- T.C. Berlin Büyükelçiliği (2002). Türklerin Almanya'ya Uyumları Bazı Klişeler ve Araştırma Sonuçları [Adaptation of Turkish People to Germany, Some Clichés

and Research Results]. Retrieved December 2007 from www.tcberlinbe.de/tr/arsiv/2002.

II. . Uluslar Arası Eğitim ve Kültür Bağlamında Avrupalı Türkler Kongresi (2009). [II. International Congress of European Turks within the context of Education and Culture] Final Declaration Retrieved Jun 2009 from <http://www.binifikir.be/news/141/ARTICLE/4484/2009-05-16.html>.

Yurt Dışında Yaşamış Üniversite Öğrencilerinin İyi-Oluş Düzeyleri: Hacettepe Üniversitesi Örneği

(Özet)

Problem Durumu: İnsan var olduğu andan itibaren içinde bulunduğu çevreye uyum sağlamaya çalışır. Uyum çabaları beraberinde bir takım güçlüklerle de neden olabilmektedir. Bu nedenle insanı anlama ve uyumunu artırmaya yönelik çabalar daha çok yaşanan sorunlara odaklanmıştır. Bununla birlikte 1990'lardan itibaren, psikoloji alanında yapılan araştırmaların sadece olumsuz yaşantıları anlamaya ya da bunların etkilerini ortadan kaldırmaya yönelik olmadığını söylemek mümkündür. Son on yılda mutluluk (Happiness), iyilik hali (Wellness), iyi oluş (Well-being) gibi kavramlar psikolojinin inceleme konusu olmuştur. Mutluluk, iyilik hali, iyi-oluş, kavramların bir takım değişkenlerle ilişkisi araştırılmış ya da bireylerde pozitif özelliklerin artırılmasına yönelik psiko-eğitim programları geliştirilmiş ve programların etkililiği araştırılmıştır.

İyi-oluş kavramı, genel olarak bilişsel ve duyuşsal olmak üzere iki boyutta incelenmiştir. Duyuşsal boyut, olumlu ve olumsuz duygulara odaklanırken, bilişsel boyut daha çok genel ve özel yaşam alanlarına ilişkin doyumla açıklanmıştır. Bir başka deyişle, iyi-oluş kavramı, insanların olumlu ve olumsuz duygusal durumlarını, yaşam doyumlarını ve özel alanlarındaki doyumlarını içeren geniş bir kavramdır. İyi-oluş kavramını açıklamaya yönelik, altı bileşenden oluşan teorik bir model önerilmiştir. Bu bileşenler; *kendini kabul, başkalarıyla olumlu ilişkiler kurma, otonomi, çevreyi kontrol etme, yaşama amacı ve kişisel gelişimdir*. Bu çok boyutlu model, olumlu fonksiyonları, Erikson'un psikososyal gelişim dönemleri ve Maslow'un kendini gerçekleştirme kavramı ile bütünleştirmiştir. Bu model iyi-oluş kavramının teorik temel eksikliğine tepki olarak ortaya konmuştur. Birbirleriyle orta derecede ilişkili altı yapı tanımlanmış olmasına rağmen, birbirinden yeterince bağımsız ve ayrı yapılar olarak görülmüşlerdir.

Avrupa'ya göç eden Türk Ailelerinin üçüncü kuşak gençlerinin içinde yaşadıkları kültüre uyum sağlamış görünmelerine karşın, üstesinden gelmeleri gereken sorunlar halen devam etmektedir. Sekizinci Beş yıllık kalkınma planı, Özel İhtisas Komisyonu Raporu'na göre "eğitimde eşitsizlik" süregelen önemli sorunlardan bir tanesidir Belçika, Almanya gibi çok sayıda Türkün yaşadığı Avrupa Ülkelerinde, o ülkelerin kendi gençleri ile karşılaştırıldığında yüksek öğretime devam edebilen Türk öğrencilerinin oranlarının çok düşük olduğu dikkati çekmektedir.

Yaşamlarının büyük bir bölümünü Avrupa ülkelerinde geçirdikten sonra, yüksek öğrenimleri için Türkiye'ye dönen gençlerin üniversite yaşamları sırasında üstesinden gelmeleri gereken sorunların, yaşamlarını Türkiye'de geçiren akranlarına oranla çok daha karmaşık olduğu düşünülmektedir. Yurt dışında yaşamış ve Türkiye'de yüksek öğrenimlerini sürdüren gençlerinin profillerini ortaya koyan bir çalışmada, söz konusu gençlerin Türkiye'de doğup büyüyen yaşlılarına oranla, ortaöğretime göre üniversitede ders çalışmaya daha çok vakit ayırdıkları, daha az sosyal etkinliğe katıldıkları saptanmıştır. Söz konusu çalışmadan yola çıkarak, Avrupa'da yaşamış ve yüksek öğrenimlerine Türkiye'de devam eden gençlerin iyi oluş düzeylerinin incelenmesine karar verilmiştir.

Araştırmanın Amacı: Bu çalışmada, yurt dışında en az bir yıl yaşamış ve yüksek öğrenimlerini Türkiye'de sürdüren üniversite öğrencilerinin, "cinsiyet", "yurt dışında yaşama süresi", "okul başarısı", "ailenin ekonomik durumu", "harçlık", "arkadaş ilişkileri", "öğretim elemanları ile ilişkiler" ve "algılanan ana-baba tutumuna" göre iyi-oluş düzeylerinin değişip değişmediğini ortaya koymak amaçlanmıştır.

Yöntem: Araştırma kapsamına, 2007-2008 öğretim yılında Hacettepe Üniversitesi öğrencisi olup, daha önce en az bir yıl yurt dışında yaşamış olan öğrencilerden; Almanca, İngilizce ve Fransızca Öğretmenliği Anabilim dalında öğrenim gören öğrenciler alınmıştır. Söz konusu bölümler, yurt dışında yaşamış öğrencilerin en yoğunlukta bulunduğu bölümlerdir. En az 1-3 yıl arası sürekli yurt dışında yaşayan öğrenciler "yurt dışı yaşantısı olan öğrenciler" olarak kabul edilmiştir. Araştırma grubunda 35 erkek (%31,5), 76 kız (%68,5) öğrenci bulunmaktadır. Böylece, toplam 111 öğrenci araştırma grubunu oluşturmuştur. Öğrencilerin yaklaşık %96'sı Almanca Öğretmenliğinde okumaktadır. Öğrencilerin, yaklaşık % 64'ü ise 10 yıl ve üzerinde yurt dışında yaşamışlardır.

Öğrencilerin iyi-oluş düzeylerine ilişkin veriler "Öznel İyi Oluş Ölçeği" ile elde edilmiştir. Testin geçerlik çalışmaları kapsamında, testin ayrıcılığını test etmek için alt-üst %27'lik puan ortalamaları t testi ile karşılaştırılmış ve sonuç anlamlı bulunmuştur (t:26.88, p<.01). Ayrıca paralel form geçerliği için, Beck Depresyon Envanteri kullanılmış. Her iki ölçekten alınan puanlar arasında. 01 düzeyinde -.70 değerinde zıt yönlü bir ilişki katsayısı bulunmuştur. Ölçeğin Cronbach Alfa güvenilirlik katsayısı. 93, test tekrar test güvenilirlik katsayısı ise. 86'dır. Verilerin analizinde, iki grup ortalaması arasındaki farkı test etmek için, bağımsız örneklem grupları için kullanılan t testi, ikiden fazla grup ortalaması arasındaki farkı test etmek için tek yönlü varyans analizi (ANOVA) kullanılmıştır. Gruplar arası farkın kaynağını tespit etmek için LSD ve Tukey çoklu karşılaştırma testleri kullanılmıştır.

Araştırmanın Bulguları: Verilerin analizi sonucunda elde edilen bulgulara göre, yurt dışında yaşamış üniversite öğrencilerinin; cinsiyet, yurt dışında kalma süresi, ailenin ekonomik durumu ve yüksek öğretimdeki başarı algılarına göre iyi-oluş puanları arasında anlamlı bir fark bulunmamaktadır. Bununla birlikte; Yurt dışında yaşamış üniversite

öğrencilerinin, günlük harçlık miktarlarına, üniversitedeki arkadaş ilişkilerine, üniversitedeki öğretim elemanları ile ilişkilerine ve algıladıkları anne-baba tutumlarına göre iyi-oluş puanları arasında anlamlı bir fark bulunmuştur. Farkın kaynağı araştırıldığında, harçlık miktarını *yeterli* görenlerle, *yetersiz* görenlerin, üniversitedeki arkadaş ilişkilerinin *oldukça iyi* olduğunu ifade edenlerle, *iyi* olarak ifade edenlerin, üniversite öğretim elemanları ile ilişkilerinin *oldukça iyi* olduğunu ifade edenlerle, *hiç iyi* olmadığını ifade edenlerin, yine üniversitedeki öğretim elemanları ile ilişkilerini *kısmen iyi* olarak ifade edenlerle, *hiç iyi* olmadığını ifade edenlerin ve ana-baba tutumlarının *demokratik ve hoşgörülü* olduğunu ifade edenlerle, *ilgisiz* olduğunu ifade edenlerin, iyi oluş puanları arasında anlamlı bir fark gözlenmiştir.

Araştırmanın Sonuçları ve Öneriler: Araştırmada, yurt dışında yaşamış üniversite öğrencilerinin iyi oluş düzeylerinin cinsiyete, yurt dışında kalma sürelerine, yüksek öğretimdeki başarı algılarına ve ailelerinin ekonomik durumlarına ilişkin algılarına göre değişmediği saptanmıştır. Bununla birlikte, harçlıklarının “yeterli” olduğunu düşünen öğrencilerin iyi oluş düzeylerinin, “yetersiz” olduğunu düşünenlere göre, üniversitedeki arkadaş ilişkilerini “oldukça iyi” olarak tanımlayan öğrencilerin iyi oluş düzeylerinin, “iyi” olarak tanımlayanlara göre daha yüksek olduğu görülmüştür.

Öğrencilerin üniversitedeki öğretim elemanları ile ilişkilerine göre öğrencilerin iyi oluş düzeyleri incelendiğinde, öğretim elemanları ile ilişkilerinin “oldukça iyi” olduğunu ifade eden öğrencilerin iyilik halleri, “hiç iyi olmadığını” ifade edenlere göre daha yüksek olduğu ortaya konmuştur. Benzer bir şekilde, üniversitedeki öğretim elemanları ile ilişkilerini “kısmen iyi” olarak tanımlayan öğrencilerin iyi oluş düzeylerinin, “hiç iyi olmadığını” tanımlayanlara göre daha yüksek olduğu görülmüştür.

Ana-baba tutumlarına göre iyi oluş düzeylerine bakıldığında, ana-baba tutumlarının “demokratik ve hoşgörülü” olduğunu ifade eden öğrencilerin iyi oluş düzeyleri, “ilgisiz” olduğunu ifade edenlere göre daha yüksek bulunmuştur.

Araştırmada elde edilen sonuçlar doğrultusunda, yurt dışında yaşamış ve Türkiye’de yüksek öğrenimlerine devam eden üniversite öğrencilerinin; kişilerarası iletişim becerilerini artırmaya yönelik hizmetlerin sunulması önerilmektedir. Bu sayede, öğretim elemanları ve arkadaşlarıyla ilişkilerini geliştirmelerine yardım edilmelidir. Ayrıca, öğrencilerin harçlıklarının ihtiyaçlarını karşılayacak düzeye getirmek için düzenlemeler yapılmalı, burs olanakları artırılmalıdır. Anne-babalara çocukları hangi yaşta olursa olsun, demokratik ve hoşgörülü davranmaları yönünde eğitimler verilmelidir. Benzer gruplarla çalışacak araştırmacılara, öğrencilerin Türkiye’ye geldikleri ilk yılda araştırma kapsamına alınmaları önerilmektedir.

Anahtar Sözcükler: İyi-oluş, üniversite öğrencileri, yurt dışında yaşama, ana-baba tutumu, kişilerarası ilişkiler.

Effects of Reflective Writing in Mathematics Methods Courses on Pre-Service Teachers' Perceptions

Renan Sezer*

Suggested Citation:

Sezer, R. (2010). Effects of reflective writing in mathematics methods courses on pre-service teachers' perceptions. *Eğitim Araştırmaları - Eurasian Journal of Educational Research*, 40, 189-206.

Abstract

Problem Statement: The use of writing as a learning tool is well-documented in the literature. The emphasis on writing in mathematics courses has been the focus of attention in the last two decades; however, writing in mathematics courses has mostly been restricted to journal writing and perceived as an add-on by many mathematics teachers. Thus the question remains as to whether teachers' beliefs and practices towards writing as a learning tool in mathematics have changed.

Purpose of the Study: The study investigates whether incoming pre-service teachers' perceptions towards writing as a learning tool in general differs from their perceptions towards mathematics writing specifically, whether writing in mathematics methods courses has any effect on future teachers' perceptions of writing as a learning tool in mathematics, and whether integrating writing in a mathematics methods course induces pre-service teachers to incorporate writing in their own courses.

Method: Control and experimental groups were established among 166 incoming education students at an urban college. During the 12-week semester, the experimental group had 11 writing assignments and a term project in which they wrote a lesson plan, while the control group had only the latter. Pre- and post-course surveys were conducted to gauge perceptions of writing as a learning tool generally, and in mathematics courses particularly, including in elementary courses. Differences in perceptions were computed and compared between control and

* Associate. Prof. Dr., City University of New York, Faculty of Mathematics, rsezer@lagcc.cuny.edu

experimental groups. The lesson plans from both groups were collected and their inclusion of writing-to-learn activities assessed.

Results: Significant differences were found between incoming teacher candidates' perceptions of writing-to-learn activities generally and in mathematics courses specifically. Significant differences were also observed in the pre- and post-course comparisons of the Perceived Benefits of Writing-to-Learn Activities in Mathematics between the control and experimental groups, in favor of the latter. No significant difference was observed between the two groups' pre- and post-course comparisons regarding the Perceived Benefits of Writing-to-Learn Activities in Mathematics for Elementary School Students survey; however, significant differences were observed in their implementation of writing-to-learn activities in their lesson plans.

Recommendations: Theoretical pedagogical knowledge is not sufficient for pre-service teachers to incorporate writing into their own lessons. Though teacher candidates may learn by rote that writing-to-learn activities are useful, only pre-service modeling ensures that their teaching will reflect their knowledge in the classroom.

Keywords: Writing-to-learn, writing-intensive, mathematics, teacher training, pre-service

Mathematics is often considered an abstract and formidable subject by non-majors (Ferguson, 2005). This stems mostly from the way mathematics is taught. Brownell (1987) argues that majority of mathematics instruction consists of a teacher giving an algorithm that the student is then asked to practice. Mathematics education has thus been reduced to the manipulation of symbols, and is limited to the algorithmic solutions of routine problems that do not require creativity (Borasi & Rose, 1989). Borasi and Rose (1989) argue that while this educational approach can bring short-term and partial success to some students, it neither supports deep understanding nor develops a strong knowledge-base; moreover, it adversely affects students' attitudes. Conversely, among different approaches to improve mathematics learning, one that is promising is integrating reflective writing into mathematics courses (Connolly & Vilaridi, 1989). Thus, in the last two decades, increased attempts to integrate writing and critical thinking into our teaching in general and into mathematics teaching in particular have been made (Lipman, 1988; Norris, 1985; McPeck, 1981; Siegel, 1980; NCTM, 2004; Silver, 1999).

The National Council of Teachers of Mathematics (2004) has asserted that writing in mathematics courses not only gives students a chance to communicate mathematical ideas but that doing so deepens conceptual understanding. Writing in one's own words helps internalize complex concepts and explaining these concepts in layman's terms indicates true understanding (Beins, 1993). Dunn (2000) surmises, "the best way to understand something is to write about it" (p. 130). The National Council of Teachers of Mathematics encourages writing in mathematics classes at all

levels, not only because it enables students to speak the language of mathematics more fluently, but also because it promotes active conceptual learning. Not only is this true in K-12 mathematics education, but it has particular benefits in the mathematics pedagogy courses as well.

Further studies have concluded that writing in mathematics courses not only improves students' attitudes toward mathematics but also has a positive effect on mathematics performance (Abdalkhani & Menon, 1998; Wood, 1992). McGovern and Hogshead (1990) summarize the learning objectives of writing assignments in four categories: promotion of student learning, enhancement of both critical thinking and reflection skills, development of writing skills, and assessment. Writing helps to increase both students' ability to communicate mathematically and their appreciation for the subject and its value (Davison & Pearce, 1988); moreover, improvement in students' writing skills can be tangibly observed (Abel & Abel 1988). While McGovern and Hogshead (1990), in their four categories, note the benefit of writing in student assessment, they omit the valuable information teachers thus obtain about student understanding. This information is a useful tool in instructional decisions and curriculum assessment (NCTM, 2004; Nahrgang & Petersen, 1986; Angelo, 1995). Angelo (1995) suggests that faculty should frequently and informally collect information on classroom assessment and use this information to make both their teaching and students' learning more effective. Teachers can start the next class by responding to the questions themselves and thus modeling the thinking skills they were hoping for among their students. For college students, an interactive full-class discussion of assignments and students' responses is most effective (Angelo, 1995). In short, writing not only helps to develop student understanding and make student assessment more accurate, but it also helps to create interactive classrooms and more responsive mathematics curricula.

Quinn and Wilson's (1997) indicate that although teachers' beliefs are changing regarding writing pedagogy in mathematics classes, this belief is not reflected in their teaching practices. Moreover, the majority of mathematical writing done seems to be limited to explanation of solutions and journal writing (Quinn & others, 1997; Silver, 1999; O'Shea, 2006). Research also shows that as the grade levels increase, mathematics teachers' regard for writing in mathematics classes decreases; in higher grades teachers feel that it improved writing skills more than mathematical understanding and see writing as non-essential in mathematics (Quinn & others, 1997; Silver, 1999). These beliefs, coupled with time constraints (to fulfill the syllabus, to respond to assignments) and a perception of students' poor writing skills, deterred teachers from using writing assignments in their mathematics classes (Quinn & others, 1997; Silver, 1999). Abel and Abel (1988), as well as Dunn (2000), suggest one solution that can both promote interactive learning and save time: teachers need not have the sole responsibility of giving feedback; peer responses to writing are also beneficial.

Unfortunately, most pre-service teachers have themselves learned mathematics in a traditional atmosphere where learning is by rote; thus, they possess neither a strong mathematics background, nor the conceptual maturity necessary for a deeper understanding of the subject matter (Kalder, 2007; McDiarmid & Wilson, 1991). Though these pre-service teachers often may appear to function well by solving

problems whose solutions can be formulated algorithmically, and thus getting the “right answer”, future elementary school teachers need the ability to explain underlying concepts. Writing in pedagogy courses can be used as a tool for both learning mathematics concepts and clarifying future teachers’ thinking. This can move the pre-service teacher beyond performing algorithms to a coherent conceptual understanding with the language to explain it. It gives future teachers an improved ability to express their thought process in a logical and coherent manner.

The use of writing in pedagogy classes can also serve to immerse future teachers in an ever-evolving discussion of best practices in the mathematics classroom. In recent decades, there have been differences of opinion regarding how mathematics should be taught. While reform in mathematics education has been encouraged by some (NCTM, 2004; AMATYC, 2006), this initiative has been resisted by others, such as the organization HOLD (Honest Open Logical Decisions on Mathematics Education Reform). In a methods class, writing can be used to discuss alternative teaching methods and assessment strategies. Often, in-coming pre-service teachers who are not familiar with mathematics reform voice the traditional views with which they have been educated. Writing early and often in their pedagogical education can help future teachers become aware of these assumptions and to challenge them. These initial writings can later be used to assess the learning of mathematics-education students in both pedagogical approaches and writing skills.

In order to improve the future elementary school teachers’ understanding of mathematics and their ability to convey this understanding, a writing-intensive mathematics methods course was developed for pre-service elementary school teachers at an urban college. This was the first mathematics methods course in a series of two. In the United States, some elementary schools cover grades 1 through 5, while others include the 6th grade. Thus topics ranged from 1st to 6th grade, beginning with the whole numbers and progressing to percents. During the 12-week instructional period, students were given 11 writing assignments.

The first writing assignment asked students to write how they envisioned their role as an elementary school mathematics teacher. What kind of classroom atmosphere would they provide? what kind of strategies would they use? how would they teach; what evaluation tools would they use? etc. This first assignment was given with the purpose of gauging students’ perceptions of teaching entering the class, and creating a window onto changes occurring over the semester.

The pre-service teachers had difficulty fathoming the challenge of a first-grader when confronted with addition operations that involved “carrying over” or subtraction operations that required “borrowing.” Due to years of familiarity with these concepts, base-ten blocks, used in the explanation of these ideas, seemed unnecessary and redundant. In order to take the pre-service teachers out of their comfort zone, the second writing assignment asked them to explore how many different numbers could be written in base 7 and base 12 respectively, using a single digit. It asked them to explain how addition is done in base 7 and 12 by carrying over and then to generalize this concept to any addition operation performed in base n . Furthermore, students were expected to explain the concept of “borrowing” in subtraction between two numbers written in base 8, base 9 and base n . While working on this homework, they realized they did not have as thorough an

understanding of how numbers were written and the underlying mathematics involved in performing addition/subtraction operations as they had previously thought.

The third assignment asked them to analyze the division algorithm. Long division is the only operation among the basic four operations that starts with the biggest digit, i.e. the leftmost digit of the dividend. In this assignment, students were asked to explore whether they could start the division algorithm from the rightmost digit (ones digit) and were to answer questions based on this shift: would they get the right answer? would their record-keeping strategy change? what, if any, were the advantages of the traditional algorithm? The aim of this activity was to have students think of non-routine approaches, look at basic mathematics through a different lens and recognize that it is actually imbedded with novel ideas.

The fourth writing assignment questioned why a common denominator is necessary when adding and subtracting fractions, but is not required in multiplication. The fifth writing assignment focused on fraction division and questioned the "flip the second number and multiply" rhetoric, whereas the sixth assignment asked them to explore the reason two equivalent ratios would give the same product when "cross-multiplied." The goal of these assignments was to encourage students' reflection on mathematical reasoning and emphasize the logic behind mathematical principles; thus, students became divorced from the perception that mathematics is a collection of isolated "rules" that need to be followed without comprehension. The emphasis was shifted from a procedural perspective to a conceptual one, echoing the research of Ma (1999): "From a procedural perspective, arithmetic algorithms have little or no connection with other topics and, are isolated from one another" (118). Asking the "why" questions is the first step to conceptual understanding in mathematics.

The seventh assignment related to percentages. Students considered the following question: "If an amount is raised by a percentage and this new amount is later reduced by the same percentage, is the original amount obtained? Develop hints/ guided questions to help a confused elementary school student with his/her thought process." For many elementary school students -- and some adults -- who only perceive the situation as forming two inverse operations, the result is baffling. It is therefore necessary that the elementary school teachers predict where the difficulty lies and develop tools that lead the student by well-placed hints, without giving the answers.

The eighth assignment involved problem solving and focused on expanding the Polya's steps to include writing a synthesis of the problem solved, independent of the parameters of the given problem, which can change. It is the author's perception that while many problems of the same category, with varying numbers, are solved in the classroom and in the homework, the overarching idea utilized in its solution is seldom verbalized.

The ninth assignment was an open-ended one, asking students to choose a mathematics concept and investigate how they could make interconnections between this concept and other mathematical areas and principles taught in the same grade level, as well as how connections could be formed between this chosen concept and

other, more advanced, mathematical concepts taught in other grade levels. The aim of this activity was to foster a deeper and broader understanding between mathematical concepts, weaving them together into a coherent whole, thus forming a cohesive body of knowledge, rather than learning fragmented and isolated facts. This assignment also hoped to illustrate that elementary school mathematics focuses on basic, yet fundamental, mathematical notions.

The tenth assignment directed the student to various Internet sites featuring different views of mathematics education: some congruent with NCTM Standards, others supporting a more traditional approach. The curriculum of this methods course included the NCTM Standards; however, to express only one perspective in a methods course meant limiting students' visions and reflections, and would lead to a knowledge devoid of critical thinking - precisely what we aimed to avoid. Reading contradictory opinions on mathematics education, combined with their own experiences as students, encouraged pre-service teachers to weigh their options and develop their own unique outlook. This assignment also enabled the instructor to "hear" individual student views - something limited class time would not easily accommodate. The eleventh assignment was a duplicate of the first assignment and was given to see if and how students' perceptions had changed: perceptions of what it means to teach and learn mathematics, of the activities used to stimulate the process, and of the most effective evaluation criteria.

Research Questions

This paper aims to answer the following research questions:

- 1) Are the incoming pre-service teachers' perceptions towards writing, as a learning tool in general, different than their perceptions towards mathematics writing in particular?
- 2) What effect, if any, does writing in mathematics methods courses have on future teachers' perceptions of writing as a learning tool in mathematics?
- 3) What effect, if any, does integrating writing in a mathematics methods course have on pre-service teachers' incorporation of writing assignments in their own teaching practices?

Method

Two writing-intensive sections of the first mathematics methods course in a series of two were chosen as the experimental group. There were multiple sections of the course offered in any given semester. The students registering to the writing-intensive section did not know ahead of time that this teaching methodology would be utilized; thus, self-selection was avoided. The control group was chosen from the three other sections, based on the instructors' willingness to devote class time to pre- and post-course surveys. The two experimental groups had a total of 74 students, while the three control groups had a total of 88 students.

To answer the first question of the study, a subset of Writing-to-Learn Attitude Survey questions were utilized, specifically Perceived Benefits of Writing-To-Learn Activities (Table 1). A similar Perceived Benefits of Writing-To-Learn Activities in

Mathematics Survey (Table 2) was then created by altering each question in the original survey, to focus on writing in mathematics courses in particular. The survey questions were answered on a 5-point Likert scale, ranging from 1 for 'strongly disagree' to 5 for 'strongly agree.' After reverse scoring of the negatively worded items, Cronbach's alpha was calculated and the reliability coefficient was found to be .90 (Schmidt, 2004).

For the first question only, the control and experimental group was regarded collectively and a pre and post design was not conducted. The perceived benefits of writing were compared to the perceived benefits of writing in mathematics classes in particular. The Perceived Benefits of Writing-To-Learn Activities was administered to students in the first day of the semester, followed by Perceived Benefits of Writing-To-Learn Activities in Mathematics Survey, given in the second class session. The questionnaires were administered anonymously to both groups; however, to enable pre- and post-pairings, students identified themselves by the last four digits of their phone numbers. Responses to corresponding questions in the two surveys were then compared using a paired- samples t-test, utilizing the SPSS software.

To answer the second question of this study, two tools were utilized:

1) Perceived Benefits of Writing-To-Learn Activities in Mathematics; 2) Perceived Benefits of Writing-To-Learn Activities in Mathematics for Elementary School Students.

The Perceived Benefits of Writing-To-Learn Activities in Mathematics was designed as indicated above. Using this questionnaire, pre and post data were collected from experimental and control groups, in the first and last week of the semester, respectively. The difference in the pre and post data scores of the Perceived Benefits of Writing-To-Learn Activities in Mathematics for the control and experimental groups were compared, using the independent-samples t-test. The t-test was performed utilizing the SPSS software package.

Secondly, Perceived Benefits of Writing-To-Learn Activities in Mathematics for Elementary School Students (Appendix A) was created, using 10 of the 11 questions in the Perceived Benefits of Writing-To-Learn Activities in Mathematics Survey. The questions were phrased to target elementary school students' learning of mathematics through writing, rather than that of the person surveyed. The one question that was omitted from the survey was "Writing to different audiences makes me aware of how much the reader or listener affects the way I state information and concepts." This question was disregarded since it is not applicable to elementary school students. This survey was administered to pre-service teachers in the control and experimental groups, in the first and last weeks of the semester. The difference in the pre- and post-data scores of the Perceived Benefits of Writing-To-Learn Activities in Mathematics for Elementary School Students was computed. These differences for the control and experimental groups were compared using independent- samples t-test. The t-test was performed utilizing the SPSS software package.

The term project assigned to all sections of this course was utilized to answer the third question of the study. The first part of the project asked pre-service teachers to

visit an elementary mathematics class and write their observations, along with their critique of the lesson. The second part asked them to write the next lesson plan for the class they visited, based on what was covered in the class they observed and what should follow. This lesson plan included activities that would take place in class; materials that would be used; and forethought about what could possibly go wrong, while including outlines of strategies or activities specifying what to do in these cases, and homework to be assigned following the lesson. The study also compared the number of pre-service teachers who integrated writing assignments in the second part of their project between the experimental and control groups.

Results

Regarding the First Question of the Study

The descriptive statistics regarding the Perceived Benefits of Writing-To-Learn Activities are presented in Table 1. Table 1 indicates that the highest mean score among the items corresponds to item 9, "Brainstorming, freewriting, or listing ideas before writing helps me find out what I know and think about a topic," with a mean score of 4.06. The lowest mean score among the items corresponds to item 25, "The technical aspects of writing (punctuation, spelling, etc.) are more important than other aspects (concept formation, clarity, etc.)," with a mean score of 2.62.

The descriptive statistics regarding the Perceived Benefits of Writing-To-Learn Activities in Mathematics are presented in Table 2. Table 2 indicates that the highest mean score among the items corresponds to item 22, "Writing to different audiences makes me aware of how much the reader or listener affects the way I state information and concepts," with a mean score of 3.80. The lowest mean score among the items corresponds to item 30, "I use journals to enhance my understanding of course materials," with a mean score of 2.51.

The results of the paired-samples t-test, comparing the data on the Perceived Benefits of Writing-To-Learn Activities questionnaire to that of the Perceived Benefits of Writing-To-Learn Activities in Mathematics is given in Table 3. As a result of the analysis, the mean of the Perceived Benefits of Writing-To-Learn Activities was found to be ($\bar{X} = 38.99$) and the mean of the Perceived Benefits of Writing-To-Learn Activities in Mathematics was found to be ($\bar{X} = 34.99$). The paired-samples t-test indicates that there are significant differences between the perceived benefits of writing-to-learn activities and the use of such activities in mathematics courses ($t_{(161)}=14.252$; $p < .05$). While significant differences were observed for items 2, 9, 11, 19, 20, 30, differences observed for items 7, 17, 22, 25, 28 were not significant. The items that had the highest differences were item 20 (difference of 1.14) and item 11 (difference of 0.74). The items with least differences were 22, 28 (both with difference of 0.04) and item 25 (difference of 0.06).

Table 1*The Descriptive Statistics of Perceived Benefits of Writing-To-Learn Activities*

Item No.	Item	\bar{X}	sd
2	Impromptu focused writing in class helps me solve problems or clarify concepts.	3.57	.72
7	Informal notes and letters to classmates about course material help me understand difficult material.	3.20	.89
9	Brainstorming, freewriting, or listing ideas before writing helps me find out what I know and think about a topic.	4.06	.83
11	Admit slips [Admit slips are brief written responses to a question about the planned topic for a class] make it easier to begin thinking about what will be covered in a class.	3.82	.80
17	Writing microthemes (i.e., brief summaries) makes me aware of the most important points in reading assignments.	3.84	.80
19	Critiquing a classmate's writing for conceptual clarity results in increased understanding for both of us.	3.59	.88
20	Writing personal experience pieces makes me see connections between what I am learning and my own life.	3.81	.73
22	Writing to different audiences makes me aware of how much the reader or listener affects the way I state information and concepts.	3.84	.76
25	The technical aspects of writing (punctuation, spelling, etc.) are more important than other aspects (concept formation, clarity, etc.).	2.62	.77
28	Exit slips [Exit slips are brief written responses to a question about the topic covered in a class] help me remember the important points covered in class.	3.63	.67
30	I use journals to enhance my understanding of course materials.	3.01	.95
Total		38.99	3.43

Table 2

The Descriptive Statistics of Perceived Benefits of Writing-To-Learn Activities in Mathematics Courses

Item No.	Item	\bar{X}	sd
2	Impromptu focused writing in class helps me solve problems or clarify concepts.	3.31	.84
7	Informal notes and letters to classmates about course material help me understand difficult material.	3.09	.78
9	Brainstorming, freewriting, or listing ideas before writing helps me find out what I know and think about a topic.	3.65	.61
11	Admit slips [Admit slips are brief written responses to a question about the planned topic for a class] make it easier to begin thinking about what will be covered in a class.	3.08	.83
17	Writing microthemes (i.e., brief summaries) makes me aware of the most important points in reading assignments.	3.73	.72
19	Critiquing a classmate's writing for conceptual clarity results in increased understanding for both of us.	3.01	.90
20	Writing personal experience pieces makes me see connections between what I am learning and my own life.	2.67	.77
22	Writing to different audiences makes me aware of how much the reader or listener affects the way I state information and concepts.	3.80	.80
25	The technical aspects of writing (punctuation, spelling, etc.) are more important than other aspects (concept formation, clarity, etc.).	2.56	.74
28	Exit slips [Exit slips are brief written responses to a question about the topic covered in a class] help me remember the important points covered in class.	3.59	.62
30	I use journals to enhance my understanding of course materials.	2.51	.82
Total		34.99	2.97

Table 3

The Descriptive Statistics and Paired t-test Results of the Perceived Benefits of Writing-To-Learn Activities (PBWLA) and the Perceived Benefits of Writing-To-Learn Activities in Mathematics Courses (PBWLAM)

Group	N	\bar{X}	SD	DF	t	p
PBWLA	162	38.99	3.43	161	14.252	.000
PBWLAM	162	34.99	2.97			

Regarding the Second Question of the Study:

The result of the independent-samples t-test, comparing the difference of the pre and post surveys of the Perceived Benefits of Writing-To-Learn Activities in Mathematics between control and experimental groups, is given in Table 4. As a result of the analysis of the mean difference of the pre and post surveys for the Perceived Benefits of Writing-To-Learn Activities in Mathematics was found to be ($\bar{X} = 1.14$) for the control group and ($\bar{X} = 4.08$) for the experimental. The independent-samples t-test indicates that there is significant difference between the

control and experimental groups ($t_{(113.038)} = -10.120$; $p < 0.05$). The effect size $\eta^2 (r^2)$, can be calculated, using the formula,

$$\eta^2 = \frac{t^2}{t^2 + (n_1 + n_2 - 2)} = \frac{(-10.120)^2}{(-10.120)^2 + (88 + 74 - 2)} = 0.39$$

Since $\eta^2 > 0.14$, we conclude that the effect size is large.

Table 4

The Independent-Samples t-test Results of the Difference of Pre and Post Survey Results of Perceived Benefits of Writing-To-Learn Activities in Mathematics Survey Between the Control and Experimental Groups

Group	N	\bar{X}	SD	DF	t	p
Control	88	1.14	1.29	113.038	-10.120	.000
Experimental	74	4.08	2.21			

The result of the independent-samples t-test, comparing the difference of the pre- and post-course surveys of the Perceived Benefits of Writing-To-Learn Activities in

Mathematics for Elementary School Students, between control and experimental groups, is given in Table 5.

As a result of the analysis of the mean difference of the pre- and post-course surveys, the Perceived Benefits of Writing-To-Learn Activities in Mathematics for Elementary School Students was found to be ($\bar{X} = 5.15$) for the control group and the ($\bar{X} = 5.64$) for the experimental group. The independent-samples t-test indicates that the difference between the control and experimental groups was not significant ($t_{(134.943)} = -1.454$; $p > .05$).

Table 5

The Independent Samples t-test Results of the Difference of Pre and Post Survey Results of Perceived Benefits of Writing-To-Learn Activities in Mathematics for Elementary School Students Survey Between the Control and Experimental Groups

Group	N	\bar{X}	SD	DF	t	p
Control	88	5.15	1.80			
				134.943	-1.454	.149
Experimental	74	5.64	2.36			

Regarding the Third Question of the Study

Of the 88 students in the control group, 6 included writing-to-learn activities in the lesson plan they had prepared, while 69 of the 74 students in the experimental group had. Since the number of successes and the number of failures are both at least 5 for the control and the experimental group, the conditions for the ratio test are

satisfied. To test the claim that experimental group proportion, $\hat{p}_1 = \frac{69}{74}$ is greater

than the control group proportion $\hat{p}_2 = \frac{6}{88}$, a hypothesis test was used, utilizing the

formula
$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$
, where $\bar{p} = \frac{69 + 6}{74 + 88} = 0.4630$ and $\bar{q} = 0.5370$.

It gives us $z = 10.99$, and we conclude that with 5% significance, the experimental group's ratio of using writing assignments in their lesson plans was greater than that of the control group's.

Discussion

There is a significant difference between incoming pre-service teachers' Perceived Benefits of Writing-To-Learn Activities and Perceived Benefits of Writing-To-Learn Activities in Mathematics. Students regard the benefits of writing-to-learn activities in mathematics as less significant than in other courses. The difference is most pronounced in items 20 and 11, indicating that incoming teacher candidates have the most difficulty relating mathematics they are learning to their own lives, followed by writing admit slips reflecting on what they already know about the mathematics topic to be covered. Contrarily, there was not a significant difference in students' perception of the benefits of writing exit slips, reflecting on what they have learned at the end of class, in general, and in mathematics courses in particular. This indicates a more passive approach to mathematics learning. Therefore, a conscious effort needs to be made connecting the mathematics students' learning to their daily lives.

A writing-intensive approach to mathematics learning created a significant change in the experimental group's perception of the benefits of writing-to-learn activities in mathematics. This effect size was large ($\eta^2 = 0.39$). On the other hand, a significant difference in the Perceived Benefits of Writing-To-Learn Activities in Mathematics for Elementary Students was not observed between the control and experimental groups. This can be explained by the fact that both control and experimental groups have studied NCTM Principles and Standards and thus both discussed the communication principle, in which writing plays a vital role.

Even though teacher candidates, in both groups, indicated a belief in the benefits of writing in mathematics courses for elementary school students, when putting it into practice, a significant difference was observed between the control and experimental groups, in favor of the experimental group. This difference can be attributed to at least one of two causes: 1) Teacher candidates learn by rote that they are expected to integrate writing into their teaching practices, but they do not internalize it or believe in it; 2) Teacher candidates do not know how to integrate writing assignments into their mathematics classes.

Conclusion and Recommendations

The results of this study indicates that it is not sufficient to teach in methods courses that writing is a useful tool for learning in general and in mathematics in particular without specifically modeling how writing activities can be used. While trying to distance mathematics teaching from rote learning as much as possible, we, in turn, may fall into the same trap in our methods courses, if we only pay lip service to the pedagogies we would like future teachers to embrace. Actively engaging future teachers in writing-to-learn activities in their own mathematical explorations enables them to see the benefits of writing to their mathematics learning, while providing them with concrete examples of how they in turn can implement it in their on classes.

References

- Abel, J. P., & F. J. Abel. (1988). Writing in the mathematics classroom. *Clearing House*, 62 (4), 155-158.
- Abdalkhani, J., & Menon, R. (1998). Using writing in college mathematics courses. *Journal on Excellence in College Teaching*, 9 (3), 3-17.
- American Mathematical Association of Two Year Colleges (AMATYC). (2006). *Beyond cross roads in mathematics*. Addison-Wesley Publishing Company, Prentice Hall, Inc.: MA.
- Angelo, T. A. (1995). Beginning the dialogue: Thoughts on promoting critical thinking. *Teaching of Psychology*, 22 (1), 6-7.
- Beins, B. C. (1993). Writing assignments in statistics classes encourage students to learn interpretation. *Teaching of Psychology*, 20, 161-164.
- Brownell, W. A. (1987). Meaning and skill - maintaining the balance. *Arithmetic Teacher*, 34 (8), 18-25.
- Connolly, P., & Vilardi, T. (1989). *Writing to learn mathematics and science*. New York: Teachers College Press.
- Davison, D. M., & Pearce, D. L. (1988). Using writing activities to reinforce mathematics instruction. *Arithmetic Teacher*, 35 (8), 42-45.
- Dunn, D. S. (1996). Collaborative writing in a statistics and research methods course. *Teaching of Psychology*, 23 (38), 3.
- Dunn, D. S. (2000). Letter exchanges on statistics and research methods: writing, responding, and learning. *Teaching of Psychology*, 27 (2), 128-130.
- Dunn, D. S., & Toedter, L. J. (1991). The collaborative honors project in psychology: Enhancing student and faculty development. *Teaching of Psychology*, 18, 178-180.
- Ferguson, D. L., (2005). *Mathematical and statistical reasoning in compelling contexts: quantitative approaches for building and interrogating personal, disciplinary, interdisciplinary and worldviews*. Retrieved January 4, 2010, from http://serc.carleton.edu/files/sencer/backgrounders/sencer_mathematical-reasoning.pdf
- Lipman, M. (1988, September). Critical thinking--What can it be? *Educational Leadership*, 46 (1), 38-43.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: teachers' understanding of fundamental mathematics*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- McGovern, T. V., Hogshead, D.L. (1990). Learning about writing, thinking about teaching. *Teaching of Psychology*, 17, 5-10.
- McPeck, J. (1981). *Critical thinking and education*. Oxford: Martin Robertson,
- Nahrgang, C. L. and B. T. Petersen. (1986). Using writing to learn mathematics.

Mathematics Teacher, 79 (6), 461-465.

National Council of Teachers of Mathematics (NCTM). (2004). *Curriculum and evaluation standards for school mathematics*. Reston, Va.: NCTM.

Norris, S. P. (1985). Synthesis of research on critical thinking. *Educational Leadership*, 42 (8), 40-46.

O'Shea, J. (2006). Encouraging good mathematical writing. *International Journal of Mathematical Education in Science & Technology*, 37 (1), 101-103.

Quinn, R. J., Wilson, M. M. (1997). Writing in the mathematics classroom: Teacher beliefs and practices. *Clearing House*, 71 (1), 14-20.

Schmidt, L. A. (2004). Psychometric Evaluation of the Writing-to-Learn Attitude Survey. *Journal of Nursing Education*, 43 (10), 458-465.

Siegel, H. (1980). Critical thinking as an educational ideal. *Educational Forum* 45 (1), 7-23.

Silver, J. (1999). A survey on the use of writing-to-learn in mathematics classes. *Mathematics Teacher*, 5, 388-389.

Smith, P. C. (1995). Assessing writing and statistical competence in probability and statistics. *Teaching of Psychology*, 22 (1), 49-51.

Wood, K D. (1992). Fostering collaborative reading and writing experiences in mathematics. *Journal of Reading*, 36 (2), 96-103.

APPENDIX A: Perceived Benefits of Writing-To-Learn Activities in Mathematics Courses for Elementary School Students

2. Impromptu focused writing in mathematics class helps elementary school students solve problems or clarify concepts.

7. Informal notes and letters to classmates about mathematics material help elementary school students understand difficult material.

9. Brainstorming, freewriting, or listing ideas before writing helps elementary school students find out what they know and think about a mathematics topic.

11. A brief written response to a question about the planned topic in a mathematics class make it easier for elementary school students to begin thinking about what will be covered in mathematics class.

17. Writing microthemes (i.e., brief summaries) makes elementary school students aware of the most important points in a mathematics assignment.

19. Critiquing a classmate's mathematics writing for conceptual clarity results in increased understanding for both students.

20. Writing personal experience pieces makes elementary school students see connections between what they are learning in mathematics class and their own life.

25. The technical aspects of writing (punctuation, spelling, etc.) are more important than other aspects (concept formation, clarity, etc.) in an elementary school mathematics class.

28. Brief written responses to a question about the topic covered in a mathematics class help elementary school students remember the important points covered in mathematics class.

30. Using journals can enhance elementary school students' understanding of mathematics materials.

Matematik Metod Dersinde Yansıtıcı Yazımın Öğretmen Adaylarının Yazımın Faydalarını Algılamasına Etkisi

(Özet)

Problem Durumu: Yazımın öğrenme üzerindeki pozitif etkisi üzerine pek çok araştırma yapılmıştır. Son yıllarda, yazımın, matematik derslerinde kullanılması üzerine yoğunlaşmış, ama matematik derslerindeki yazımlar günlük tutma dışına pek çıkamamıştır. Ayrıca, matematik öğretmenlerinin yazımı hala bir ek ve yük olarak algıladıkları çalışmalarca gözlemlenmiştir. Bu durum, öğretmenlerin, matematik derslerinde, yazım tekniğini kullanmalarının faydası konusunda görüşlerinin gerçekten değişip değişmediği ve yazımı, matematik öğretmede, ne denli sınıflarında kullandıkları sorularını gündeme getirmektedir.

Araştırmanın Amacı: Bu araştırma, üç soruya yanıt vermeyi amaçlar: 1) İlkokul öğretmen adaylarının, eğitim derslerine girişte, yazımın öğrenmeye etkisi üzerine görüşleri ile matematik öğrenmeye etkisi üzerine olan görüşleri aynı mıdır?

2) Yazım ağırlıklı matematik metod dersleri, ilkokul öğretmen adaylarının yazımın matematik öğrenmedeki etkisi üzerindeki görüşlerinde anlamlı bir farklılık sağlar mı?

3) Yazım ağırlıklı matematik metod dersleri, ilkokul öğretmen adaylarının yazımı, kendi matematik derslerinde kullanımları açısından, anlamlı bir farklılık yaratır mı?

Araştırmanın Yöntemi: Araştırmanın ilk sorusunu cevaplamak amacıyla, ilkokul öğretmen adaylarının, yazımın öğrenmede ve matematiği öğrenmede etkilerini algılayışlarını araştırmak üzere, bir üniversitede verilen ilk matematik metod dersinde, 166 ilkokul öğretmen adayına, dersin ilk günlerinde, Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması ve Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması olmak üzere iki anket verilmiştir. Bu iki anketin ortalamalarının eşleştirilmiş t-testi (paired sample t-test) ile kıyaslanması, SPSS programı ile yapılmıştır. Araştırmanın ikinci sorusunu cevaplamak için iki yöntem kullanılmıştır: 1) Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması anketi, 2) İlköğretim Öğrencileri İçin Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması anketi. Rastasal (random) seçilen 166 öğrenciden 88'i kontrol grubunu, 74'ü de deney grubunu oluşturmuştur. Kontrol grubundaki öğrenciler 3 şubeden, deney grubundaki öğrenciler ise 2 şubeden gelmiş olup, her iki gruba da, dönem başında Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması anketi ve İlköğretim Öğrencileri İçin Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması anketi verilmiştir. Daha sonra, deney grubundaki öğrencilere yazım ağırlıklı matematik metod dersi uygulanırken, kontrol grubundaki öğrencilere geleneksel bir yaklaşımla matematik metod dersi verilmiştir. Deney grubundaki öğrencilere, 12 haftalık dönemde, 11 yazım ödevi verilmiş, bu ödevlerle hem matematik kavramlarını daha iyi anlamaları, hem de değişik pedagojileri tartışmaları sağlanmıştır. Dönem sonunda, kontrol ve deney grubundaki öğrencilere Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması anketi ve İlköğretim Öğrencileri İçin Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması anketleri tekrar verilmiştir. Öğrencilerin adlarını yazmadıkları ve telefon numaralarının son dört rakamı ile kodlanmış, ilk ve son anketler SPSS programı ile incelenmiştir. Kontrol ve deney grupları arasında anlamlı bir fark olup olmadığı, bağımsız gruplar için t-testi (independent samples t-test) uygulanarak saptanmıştır. Üçüncü araştırma sorusunun cevaplanmasında, her iki gruptaki öğrencilere verilen dönem sonu ödevi kullanılmıştır. Dönem sonu ödevi, her bir öğrenci için, izlediği bir matematik sınıftan yola çıkarak, bir sonraki gün için bir ders planı hazırlamayı içerir. Araştırma sorusunun cevaplanması için, kontrol ve deney grubundaki öğrencilerin, hazırladıkları ders planı ve ev ödevlerinde, matematik öğrenmeyi destekleyecek yazım ödevlerini kullanma oranları incelenmiştir. Bunun için de, deney grubundaki öğrenciler tarafından hazırlanan ders planlarında, yazım oranının daha çok olup olmadığına dair hipotez testi yapılmıştır.

Araştırmanın Bulguları: İlk sorunun bulguları: İlkokul öğretmen adaylarının Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması ve Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması arasında anlamlı bir fark olduğu bulunmuştur. Öğretmen adaylarının matematiği yazımla öğrenme etkinliklerinin katkısını algılamaları, başka konularda yazımla öğrenme etkinliklerinin faydalarını algılamalarına kıyasla çok daha düşüktür. İki anket arasındaki en büyük farklılık, öğrencilerin öğrendikleri matematikle gündelik hayatları arasında bağlantı kurmalarında gözlemlenmiştir. İkinci büyük fark ise, öğrencinin öğreneceği konu üzerinde ön bilgilerini yazmasından doğan katkı olarak ortaya çıkmıştır. Öğretmen adaylarının, başka konular ve matematik derslerinde öğrenmiş oldukları bilgiler üzerine yazmakta algıladıkları katkılar arasında istatistiksel bir fark yokken, ön bilgilerini yazmaktaki fayda algılamasında anlamlı bir farklılık ortaya çıkmıştır. Bu bulgu, öğretmen adaylarının matematik öğrenimini daha pasif algılamaları anlamına gelir.

İkinci sorunun bulguları: Matematiği Yazımla Öğrenme Etkinliklerinin Faydalarının Algılanması anketinin ön uygulaması ve son uygulaması arasındaki fark, kontrol ve deney grupları için anlamlı bir farklılık göstermiştir. Yazım ağırlıklı matematik metod dersi gören, deney grubu öğrencilerinin, ders öncesi ve sonrası, matematiği yazımla öğrenme etkinliklerinin faydalarını algılamalarında, kontrol grubuna göre büyük bir artış olmuştur. Etki düzeyi ($\eta^2 = 0.39$) olarak bulunmuştur ve bu etki düzeyi büyüktür. Fakat kontrol ve deney grubunun, ders öncesi ve sonrası, ilköğretim öğrencileri için matematiği yazımla öğrenme etkinliklerinin faydalarını algılamaları arasında istatistiksel bir fark görülmemiştir. İki grup arasındaki algılamaların arasında anlamlı bir fark olmamasının sebebi, hem kontrol hem de deney gruplarında, NCTM prensiplerinin işlenmiş olması ve bu prensiplerde yer alan iletişim prensibinde, yazımın önemli rol oynadığının vurgulanmış olması, dolayısıyla, her iki grubun da matematik öğreniminde yazıma yer verilmesi üzerine bilgilendirilmiş olmasından kaynaklanabilir.

Üçüncü sorunun bulguları: 88 kişiden oluşan kontrol grubunda 4 öğretmen adayı ders planlarında yazımla öğrenme etkinliğine yer vermişken, 77 kişiden oluşan deney grubunda ise 74 öğretmen adayı yazıma yer vermiştir. Bulgular, deney grubunda hazırlanan ders planlarında yazımla öğrenme etkinliklerine daha büyük oranda yer verildiği yönündedir.

Öneriler: Metod derslerinde, öğretmen adaylarına, yazımı öğrenmede kullanmak üzerine verilen teorik bilgiler, onların bu bilgiyi içselleştirmelerini ve uygulamalarını sağlamaya yetmemektedir. Onlara, yazımla matematik öğrenmeyi sağlayacak ödevlerin nasıl geliştirildiğini ve nasıl kullanıldığını, uygulamalı olarak göstermediğimiz sürece, bu yöntemler, ezberci bir eğitimden öteye geçmemekte ve yetiştirilen öğretmenlerin, bu teknikleri kendi sınıflarında uygulamalarını sağlamamaktadır.

Anahtar sözcükler: Yazım, matematik, öğretmen yetiştirme, metod dersleri

Validation of the Social Achievement Goal Orientation Scale in Iranian Students

Siavash Talepasand*

Fatimah Alijani**

Iman Bigdeli***

Suggested Citation:

Talepasand, S., Alijani, F., and Bigdeli, I. (2010). Validation of the social achievement goal orientation scale in Iranian students. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 40, 207-222.

Abstract

Background: A growing number of researchers have recognized the importance of studying social goals along with academic goals to better understand motivational dynamics. While the call for a Social Achievement Goal Orientation theory has received considerable discussion in Western studies, it is not clear whether the theory can also be applied to other ethnic and cultural contexts.

Aims: The objective of this study was to validate the Iranian version of the Social Achievement Goal Orientation Scale and to initially test the Social Achievement Goal Orientation theory in the context of Iranian students.

Method: A total of 403 Iranian students (159 female and 244 male) from various high schools in Saveh participated in the study. Participants completed a Social Achievement Goal Orientation Scale (SAGOS) and an Achievement Goal Questionnaire. Exploratory factor analysis, item-total correlation, and reliability analyses were undertaken to assess the psychometric properties of the SAGOS. Confirmatory factor analysis was used to provide further validation for the questionnaire.

Results: Exploratory and confirmatory factor analyses confirmed the hypothesized model of social achievement goals. The fit of the proposed three-factor model was promising. Moderate support for the three-factor structure of social goal orientation was found using scores from an abbreviated 13-item SAGOS. Convergent validity for the new measure of social achievement goals was established.

* Dr., Semnan University, Semnan- Iran, stalepasand@semnan.ac.ir

** Shahid Rajaei Teacher Training University, Tehran-Iran

*** Semnan University, Semnan-Iran

Conclusion: Our findings generally supported the trichotomous framework of the social achievement goal orientation theory with Iranian students. Consistent with the academic goal orientation theory, our findings indicated that social mastery and performance-approach and performance-avoidance goals had positive relationships with academic goals.

Keywords: Social goals; Goal orientation; Construct validity; Convergent validity; Confirmatory factor analysis

Achievement goal theory provides a context for understanding student motivation and behavior in an academic achievement setting. Achievement goals reflect how students evaluate their own competence in achievement situations, and they lead to the different ways in which students participate in and manage such situations (Pintrich & Schunk, 2002). The achievement goal framework has engendered an abundant body of literature on achievement-relevant behaviors (such as learning, interest in tasks, effort and persistence). This body of literature has distinguished two different types of goals within the achievement goal framework: mastery and performance goals. Mastery goals correspond to the desire to understand a task, acquire new knowledge, and develop abilities. Conversely, performance goals refer to the desire to show competency by trying to obtain positive judgments (Dweck, 1986; Nicholls, 1984). Recent research shows that performance goals can also be subdivided into approach and avoidance goals (Elliot, 1999), making a total of three dimensions. Elliot & Dweck (2005, p.7) have argued that there is a broader category of competence goals that are easily applied to "ordinary activities". Thus achievement goal orientation is one domain of a broader form of motivation that it is known as competence-relevant motivation. They have argued that in addition to academic setting, competence is applicable to other areas, such as social, emotional, cognitive, cultural and moral competence (p.8). Hence, competence goals can be just as present in social domains as they can be in academic domains. Recently, researchers have recognized the importance of studying social goals along with academic goals to better understand motivational dynamics (Urdan, 1997; Anderman & Anderman, 1999; Covington, 2000; Deci & Ryan, 2000; Patrick, Anderman & Ryan, 2002; Wentzel, 2000; Dowson & McInerney, 2001).

Social goal orientations are similar to academic goals in a three-dimensional framework (Ryan & Hopkins, 2003; Elliot & Church, 1997). For example, Mastery-oriented students employ intrapersonal criteria for evaluating success, "My performance is better than it was at the beginning of the semester". Likewise, a social mastery goal orientation reflects a focus on the development of competence in social relationships (Hopkins & Ryan, 2000). For example, my relationship with my friends is deeper than before.

Social performance goal orientation is counterpart with academic performance goal orientation. Students with an academic performance goal orientation focus on the demonstration of competence and use interpersonal standards to evaluate success (Ryan & Hopkins, 2003). For example, "My performance is better than other students". Similarly, a social performance goal orientation reflects a focus on the

demonstration of competence in social relationships. For example, most of all I get confirmation by others.

Just as with academic performance goals, social performance goals can be divided into approach and avoidance components. Students with a social performance-approach goal orientation focus on demonstrating behaviors that would result in positive social consequences. For example, my goal is to obtain approve by others. Students with a social performance-avoidance goal orientation focus on avoiding behaviors that would result in negative social consequences.

The objective of this study was to investigate how achievement goal theory is related to the social domain. The investigation of goal orientations in other domain may shed an additional light on competence-related motivation. The researchers chose the social domain for studying Iranian students, because Iran's education system is characterized by specific demographic, cultural and economic aspects, some of which may affect social interactions and academic goal orientations. Less focus on group work, less social interaction among students in class, large class sizes, and the weak economic status of families, teachers and schools may cause Iranian students to select different goal orientations compared with more affluent Western students.

No previous study has examined the psychometric properties of an Iranian version of the SAGOS on high school students. Therefore, this study was particularly focused on the Social Achievement Goal Orientation Scale (SAGOS) developed by Ryan to represent social achievement goal orientation, social mastery orientation, social performance-approach orientation and social performance-avoidance orientation (Hopkins & Ryan, 2000; Ryan & Hopkins, 2003).

Methods

Participants

Participants included 403 high school students from Saveh (159 females and 244 males). They were from sixth (39.5%), seventh (31.4%) and eighth (28.8%) grades. The stratified sampling method was used to select these groups.

Tools

The Social Achievement Goal Orientation Scale (SAGOS; Hopkins & Ryan, 2000; Ryan & Hopkins, 2003): The SAGOS is a 22-item measure with items written to represent three social goal orientations: social mastery, social performance-approach, and social performance-avoidance. The items are rated on a response scale of 1 (not at all true for me) to 5 (very true for me). The social mastery subscale consists of eight items with possible scores ranging from 8 to 40. Both of the social performance subscales consist of seven items with possible subscale scores ranging from 7 to 35. For every subscale, higher scores indicate a stronger endorsement of that goal.

Achievement Goals: students' achievement goals were measured using the Achievement Goal Questionnaire (AGQ; Elliot, 1999). The AGQ contains 18 items,

with six items used to assess each of the three separate goal orientations: mastery, performance-approach, and performance-avoidance. Students were instructed to simply indicate how true each item was for them using a 1 (Strongly Disagree) to 4 (Strongly Agree) scale. The possible subscale scores ranged from 6 to 24. Reliability and validity of the AGQ have been examined in numerous studies.

Procedure

The subjects attended the research and completed all questionnaires. The order of administration for the first half of the questionnaires was SAGOS and AGO. This order was reversed for the latter half of the administered questionnaires. Students were told that there were no right or wrong answers and that they could skip any questions that they did not feel comfortable answering. Students were also informed that the information in the survey would be kept confidential and that their teachers would not have access to their responses.

Results

The results of the descriptive statistics are presented in Table 1. All achievement goals have an average score lower than the midpoint of the scale. The standard deviations range from .45 to .54. The correlations of the scores range from .43 to .44 for the SAGOS and .34 to .62 for the AGQ.

Reliability

Cronbach's alpha coefficients were calculated to examine the internal consistency of test scores for each of the three social achievement goal subscales.

The analyses of the total sample ($n = 403$) yielded a Cronbach's alpha of .75, .67 and .66 for social mastery, social performance-approach, and social performance-avoidance, respectively. In social mastery, all corrected item-total correlations ranged between $r = 0.70$ ("It is important to me to have friends who truly care about me") and $r = 0.82$ ("I feel successful when I learn something new about myself and how I relate to other people "); in the social performance-approach, all corrected item-total correlations ranged between $r = 0.21$ ("I feel successful when I impress others with my personality or social skills") and $r = 0.54$ ("I want to be seen as important by other people"); and in the social performance-avoidance, all corrected item-total correlations ranged between $r = 0.32$ ("I am often concerned that others won't like me") and $r = 0.42$ ("It is important to me that I avoid looking foolish").

Table 1

Means, Standard Deviations, Coefficient Alpha Reliability Estimates, and Pearson Correlations among Achievement Variables

Variable	M	SD	1	2	3	4	5	6
SAGOS								
1.social mastery orientation	1.52 ^a	.45	.75 ^b	.44	.43	.51	.48	.34
2.social performance-approach orientation	2.06	.53		0.67	.43	.25	.39	.29
3.social performance-avoidance orientation	1.87	.53			.66	.20	.37	.37
AGQ								
4.mastery orientation	1.54	.54				.81	.62	.34
5.performance-approach orientation	1.74	.53					.74	.47
6.performance-avoidance orientation	2.02	.53						.57

Note. N = 403. All correlation was significant at the 0.01 level. SAGOS = Social Achievement Goal Orientation Scale; AGQ = Achievement Goal Questionnaire.

- a. Subscale means ranged from 1 to 5 for the SAGOS and 1 to 4 for the AGQ subscales.
- b. Main diagonal is Coefficient Alpha Reliability Estimates.

Convergent validity

As evidence of the convergent validity, Pearson correlation coefficients were computed with a measure of the Achievement Goal Questionnaire (AGQ) on the total sample. The results are displayed in Table 1. As predicted, all the correlations were positive and significant.

Exploratory factor analyses

In order to determine the number of factors and to investigate the properties of the 22 SAGOS items, an exploratory factor analysis was performed on the high school student sample. We initially conducted exploratory factor analyses of all 22 items of the SAGOS. We began with a principle component analysis to assess the number of factors in the SAGOS. Bartlett's Test of Sphericity was significant (Approx. $X^2 = 1624.55$, $df = 231$, $p < .001$), and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) = .83. In this analysis, six factors emerged with eigenvalues greater than 1.0. However, inspection of the scree plot suggested one large initial component (lambda 5.05, percentage of total variance explained 22.97%) followed by a series of smaller components (e.g., component 2 lambda 1.83, 8.3%; component 3 lambda 1.53, 6.97%; component 4 lambda 1.29, 5.9%; and etc.). Most of the SAGOS items all loaded positively on the first unrotated principle factor (loaded more than 0.24). In general, the scree plot suggested that one- to four-factor solutions were reasonable. Thus, we undertook a series of factor analyses to extract one-, two-, three-, and four-factor

solutions. We orthogonally rotated (VARIMAX) these factors to a simple structure prior to interpretation.

The first factor included most of the SAGOS items which were loaded with factor loading more than .24 on this factor. However, this factor was not clear and interpretable. For the two-factor solution, most of the highest loading items on the first factor were from the social mastery subscale (with exception of item 15). Factor one was comprised of eight of the original social mastery items, five of which were originally on the social performance-avoidance subscale (4, 5, 9, 16 and 17) and two of which were originally on the social performance-approach subscale (11 and 15). Item 2 loaded on both factors equally. There was substantially less evidence to support the first factor as reflecting the original social mastery scale. Factor two was comprised of five of the original social performance-approach items (3, 6, 10, 18 and 20) and two items that were originally on the social performance-avoidance subscale (12 and 21). Item 20 loaded on both factors equally.

For the three-factor solution, factor one still emerged as a social mastery factor, with eight of the original items loading on this factor, two of which were originally on the social performance-approach subscale (11 and 15). Item 15 loaded on both factors (.47 and .42 on factor one and two, respectively). Factor two was closer to matching the performance-avoidance orientation in this solution, with five of the original items loading on this factor. Two of the original items (12 and 21) loaded more on the performance-approach factor. Item 21 loaded on both factors (.36 and .40 on factor two and three, respectively). Factor three emerged as a performance-approach factor, with five of the original items loading on this factor. Two of the original items (11 and 15) loaded more on the social mastery factor.

The following items seemed to be problematic for their respective subscales: items 11 and 15 of the social performance-approach subscale, and items 12 and 21 of the social performance-avoidance subscale. When they were removed, a clear solution was gained. Eight of the original social mastery items clearly loaded on this factor, and five of the original social performance-approach items (3, 6, 10, 18 and 20) and five of the original social performance-avoidance items (4, 5, 9, 16 and 17) loaded clearly on their factors. There was substantially more evidence to support the three-factor solution (see Table 2). In sum, four items were removed from the original 22-item scale, resulting in a reduced 18-item version of the SAGOS [Abbreviated 18-item scale (A18IS-3FAC)]: eight mastery items, five performance-approach items, and five performance-avoidance items.

Table2*SAGOS, three-factor rotated solution with high school students*

item	Component		
	1	2	3
14	.664	.279	-.065
13	.651	.115	.143
02	.624	.090	.128
07	.606	.055	.152
19	.597	.289	-.064
22	.563	.184	-.048
01	.531	-.123	.153
08	.515	.144	.196
05	.059	.698	.007
04	.076	.637	.058
16	.231	.572	.044
17	.172	.568	.096
09	.071	.500	.118
10	.077	.051	.726
18	.096	.371	.668
06	.115	-.021	.663
03	.014	-.001	.571
20	.283	.292	.482

For the four-factor solution, factor one emerged as a personal mastery factor. Factor one was comprised of three of the original social mastery items (14, 19 and 22), one of which was originally on the social performance-avoidance subscale (17) and one of which was originally on the social performance-approach subscale (15). Factor two was closer to matching the performance-approach orientation in this solution. Factor two was comprised of five of the original social performance-approach items (3, 6, 10, 18 and 20), one of which was originally on social performance-avoidance subscale (21). Factor three appeared as a social mastery factor. Factor three was comprised of five of the original social mastery items (1, 2, 7, 8 and 13), one of which was originally on the social performance-approach subscale (11). Factor four

emerged as a performance-avoidance factor. This factor was comprised of five of the original social performance-avoidance subscale items (4, 5, 9, 12 and 16).

In the four-factor solution, the social mastery factor was divided into two subscales. The first subscale included items 8, 14, 19 and 22, which was closer to personal mastery. The second subscale included items 1, 2, 7 and 13 which was closer to social mastery. In the solution, the following items seemed to be problematic for their respective subscales: items 11 and 15 of the social performance-approach subscale, and items 17 and 21 of the social performance-avoidance subscale.

In sum, four items were removed from the original 22-item scale, resulting in a reduced 18-item version of the SAGOS [Abbreviated 18-items scale (A18IS-4FAC)]: four personal mastery items, four social mastery items, five performance-approach items, and five performance-avoidance items.

To summarize, the exploratory factor analyses indicated that a three-factor or four-factor solution would be acceptable. The four-factor solution explained more variance than the three-factor solution (48% versus 41%, respectively).

Item analysis of the three- and four- factor models

Later, we evaluated the internal consistency (coefficient alpha) and corrected the item-total correlations for the three-factor model (A18IS-3FAC): the social mastery, social performance-approach, and social performance-avoidance subscales. For the model, the coefficient alpha for the eight items on the social mastery subscale was 0.75. All the corrected item-total correlations were acceptable. All corrected item-total correlations ranged between $r = 0.33$ and $r = 0.55$. The coefficient alpha for the five items on the social performance-approach subscale was 0.66, and all corrected item-total correlations ranged between $r = 0.29$ and $r = 0.54$. The coefficient alpha for the five items on the social performance-avoidance subscale was 0.62, and all corrected item-total correlations ranged between $r = 0.31$ and $r = 0.43$.

For the four-factor model (A18IS-4FAC), the coefficient alpha for the four items on the personal mastery subscale was 0.66. All corrected item-total correlations ranged between $r = 0.33$ and $r = 0.51$. The coefficient alpha for the four items on the social mastery subscale was 0.64. All corrected item-total correlations ranged between $r = 0.39$ and $r = 0.54$.

However, it was not clear which of these two models is more justified. To assess the legitimacy of keeping these scales separate, we modeled the structure of the inter-item correlations using confirmatory factor analyses.

Confirmatory factor analyses

All CFA's were performed using LISREL 8.54 (Joreskog & Sorbom, 1993). Upon consideration of univariate and multivariate kurtosis, a maximum likelihood estimation (ML) with adjustments for non-normality was employed. Several fit indices were used to assess model fit: the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), the comparative fit

index (CFI), Goodness of Fit Index (GFI) and Adjusted Goodness of Fit Index (AGFI). It is suggested that cutoffs at or below approximately .05 for the RMSEA, at or above approximately .96 for the CFI, and at or below .07 for the SRMR indicate adequate fit (Yu & Muthen, 2002). The GFI and AGFI indices should fall between 0 and 1, with larger values indicating a better data-model fit (Joreskog & Sorbom, 1981). We used cutoffs at or above approximately .95 for the GFI and AGFI.

Assessing Model Fit: Table 3 presents the fit indices for each of the hypothesized models. The data-model fit was poor for the one-factor solution in the sample: $\chi^2 = 798.27$, $df = 209$, $P = .00$, $CFI = .66$, $SRMR = .076$, $RMSEA = .091$, $GFI = .83$, $AGFI = .80$. Inspection of the residual matrix indicated that substantial variance was not accounted for in the one-factor solution. For the two-factor solution, the data-model fit was poor, as well. For the four-factor solution, the data-model fit was presented in Table 3. Although model G is an alternate model, the current researchers do not consider the model for theoretical purposes.

Table 3

Fit statistics for the various hypothesized models (N = 403)

Model	χ^2	CFI	RMSEA	RMSEA 90%CI	SRMR	GFI	AGFI
(A) One-factor	798.27	0.66	0.091	0.084-0.097	0.076	0.83	0.80
(B) Two-factor mastery/perform	757.62	0.72	0.081	0.075-0.087	0.071	0.85	0.82
(C) Two-factor approach/avoid	819.49	0.70	0.086	0.079-0.092	0.073	0.84	0.81
(D) Three-factor mastery/approach/avoid	664.16	0.76	0.074	0.068-0.081	0.072	0.87	0.84
(E) Three-factor Mastery/approach/avoid (A18IS-3FAC)	372.47	0.84	0.067	0.059-0.075	0.057	0.91	0.88
(F) Four-factor Per mastery/Soc mastery/ approach/avoid	589.78	0.78	0.069	0.062-0.075	0.072	0.88	0.85
(G) Four-factor Per mastery/Soc mastery/ approach/avoid (A18IS-4FAC)	310.29	0.87	0.059	0.051-0.068	0.059	0.92	0.90

Note. χ^2 ML maximum likelihood; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; GFI = goodness of fit index; AGFI = adjusted goodness of fit index; Per = personal; soc = social; A18IS-3FAC = Abbreviated 18-items scale –three-factor; A18IS-4FAC = Abbreviated 18-items scale –four-factor.

The data-model fit was better for the three-factor solution than the two-factor solution, because the data-model fit for the three-factor solution mastery/approach/avoid [18-item version of the SAGOS (A18IS-3FAC)] was improved considerably. For this three-factor model (model E), $\chi^2 = 372.47$, $df = 132$, $P = .00$, $CFI = .84$, $SRMR = .057$, $RMSEA = .067$, $GFI = .91$, $AGFI = .88$.

However, the fit statistics did not suggest fit. For the model E, we inspected the instrument to identify theoretical issues associated with problematic items, and then we deleted items 8, 14, 19, and 22 from the mastery subscale. The phrase for item 14, "I like friendships that challenge me to learn something new about myself", was deleted for theoretical reasons. Probably, learning new things about oneself is related to personal mastery, not social mastery. Furthermore, the first phrase of item 19 ("I feel successful when I learn something new about myself") is also related to personal mastery, not social mastery. Upon examining the content of item 22 ("I would be successful if I had friends who accepted me for who I am"), we found that it may be representing the fear of disapproval or the feeling of being rejected, all of which are related to performance-avoidance. The content of item 8 ("It is important to me that I feel that I have friends I enjoy spending time with") may be driven by more than one latent construct, in the sense that one can "enjoy" in terms of position or nearness in a relationship.

Since the relationship between item 17 and all of the remaining performance-avoidance items was not represented well by the model, we also deleted item 17 from the performance-avoidance subscale. Finally, the proposed model gained a reduced 13-item version of the SAGOS (see Figure 1).

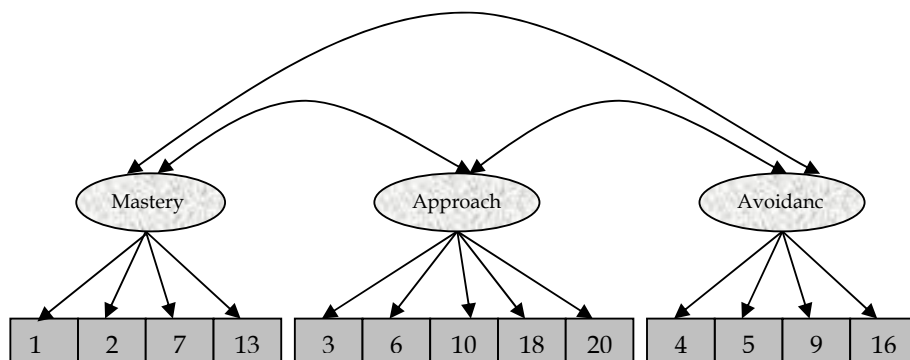


Figure 1 Final model for confirmatory factor analysis of the social achievement goal items.

Assessing Model Fit: Abbreviated 13-item SAGOS. Fit indices for the three-factor model of social goal orientation as measured by scores from the abbreviated 13-item scale indicated an acceptable overall data-model fit (see Table 4). In the model, $\chi^2 = 110.04$ and $df = 62$; thus $\chi^2 / df = 1.77$, which is lower than 2.0, and CFI = .94, SRMR = .048, RMSEA = .044, GFI = .96, AGFI = .94. Expected Cross-Validation Index (ECVI) = 0.42 [90 Percent Confidence Interval for ECVI = (0.35 - 0.50)], ECVI for Saturated Model = 0.45 and ECVI for Independence Model = 2.15. Comparing the ECVI value for the hypothesized model to the values obtained for the independence and saturated models shows that $ECVI_h < ECVI_s$, which leads to the conclusion that the hypothesized model has better predictive validity than the saturated model.

Parameter estimates. For the three-factor model, the standardized coefficients, error terms, and variance explained (R^2) were examined (see Table 4). All standardized paths were significant. However, for more than 80% of the items, less than 50% of their variance was explained by the factor for which they were written. Since, some of the standardized error terms were high, it is supposed that there is a construct - irrelevant variance (CIV). In spite of the fact that three-factor model fits, some of items had a large amount of unexplained variance.

In the social mastery subscale, item 2 has moderate and acceptable reliability ($R^2 = 0.56$), while the reliability estimates for items 1, 7 and 13 seem somewhat low (and $R^2 = 0.23$, $R^2 = 0.31$ and $R^2 = 0.27$, respectively). Here, item 2 seems to be the strongest and most reliable indicator of the latent construct social mastery orientation ($\hat{\lambda}_{21(s \text{ tan } d)} = 0.75$), followed by item 7 ($\hat{\lambda}_{31} = 0.56$), item 13 ($\hat{\lambda}_{41(s \text{ tan } d)} = 0.52$) and item 1 ($\hat{\lambda}_{11(s \text{ tan } d)} = 0.48$). In the social performance-approach subscale, item 18 has moderate and acceptable reliability ($R^2 = 0.54$), while the reliability estimates for items 3, 6, 10 and 20 seem somewhat low ($R^2 = 0.12$, $R^2 = 0.19$, $R^2 = 0.33$, and $R^2 = 0.34$, respectively). Item 18 seems to be the strongest and most reliable indicator of the latent construct social performance-approach orientation ($\hat{\lambda}_{82(s \text{ tan } d)} = 0.74$). In the social performance-avoidance subscale, item 4 has moderate and acceptable reliability ($R^2 = 0.42$), while the reliability estimates for items 5, 9 and 16 seem somewhat low ($R^2 = 0.33$, $R^2 = .18$, and $R^2 = 0.20$, respectively). Item 4 seems to be the strongest and most reliable indicator of the latent construct social performance-avoidance orientation ($\hat{\lambda}_{103(s \text{ tan } d)} = 0.64$).

Dimensionality of social and academic goal domains

Scores from the 13 SAGOS items and the 14 AGQ items (items 9, 10, 11 and 16 having been removed from AGQ) were simultaneously submitted to CFA in order to test their distinctiveness. Fit statistics for the six-factor model suggested acceptable fit (CFI = 0.90; GFI = 0.9; RMSEA = 0.048; SRMR = 0.054). These results suggest that social and academic goals are distinct and therefore can not be combined into general mastery, performance-approach, and performance-avoidance goals.

Table 4

Fit indices, standardized parameter estimates, and subscale characteristics for high school students (N=403)

Abbreviated 13-item scale							
Model	χ^2	CFI	RMSEA	RMSEA 90%CI	SRMR	GFI	AGFI
Three-factor mastery/approach/avoid	110.04	0.94	0.044	0.030-0.057	0.048	0.96	0.94
Items	Path coefficients	Error variance	R^2 value				
<i>Mastery</i>							
1	0.48	0.77	0.23				
2	0.75	0.44	0.56				
7	0.56	0.69	0.31				
13	0.52	0.73	0.27				
<i>Performance-approach</i>							
3	0.35	0.88	0.12				
6	0.43	0.81	0.19				
10	0.57	0.67	0.33				
18	0.74	0.46	0.54				
20	0.59	0.66	0.34				
<i>Performance-avoidance</i>							
4	0.64	0.58	0.42				
5	0.58	0.67	0.33				
9	0.42	0.82	0.18				
16	0.45	0.80	0.20				
	Mean	SD	<i>Mastery</i>	<i>Performance-approach</i>	<i>Performance-avoidance</i>		
<i>Mastery</i>	1.49	0.52	0.64 ^a				
<i>Performance-approach</i>	2.26	0.62	0.38 ^b	0.66			
<i>Performance-avoidance</i>	1.79	0.60	0.48	0.40	0.57		

Note. χ^2 ML maximum likelihood; CFI= comparative fit index; RMSEA = root mean square error of approximation; SRMR= standardized root mean square residual; GFI= goodness of fit index; AGFI= adjusted goodness of fit index; ^a a reliability coefficient ^b covariance

Discussion

Overall, the results of this study support the reliability, construct and convergent validity of an abbreviated 13-item SAGOS. The SAGOS was an internally consistent, multidimensional measure with a three-factor structure. Factor analyses provided support for the three-factor model. Factor one of the SAGOS suggests a social

mastery orientation that concerns a focus on developing social competence, which includes learning new behaviors, emotional growth, and improvement of relationships. Success of social mastery would be judged by whether one is improving social skills, deepening the quality of relationships, or developing one's social life in general. These findings are consistent with the hypotheses of theorists who have conceptualized social competence as social skillfulness and social abilities that promote general peer acceptance and the formation of friendships (Harter, 1982; Rubin, Coplan, Nelson, Cheah, & Lagace-Seguín, 1999).

Social goals, as well as mastery and performance goals, can be conceptualized in terms of the approach-avoidance distinction, such as the need for affiliation (approach) or the fear of rejection (avoidance) (Elliot & Thrash, 2001). The factor two of the SAGOS suggests a social performance-approach orientation. A social performance-approach goal concerns a focus on demonstrating social competence and gaining positive judgments from others indicating that one is socially desirable. Factor three of the SAGOS suggests a social performance-avoidance orientation. A social performance - avoidance goal concerns a focus on demonstrating that one does not lack social competence. The chief concern of this goal is to avoid doing something that would incur negative judgments from others and indicate social undesirability. These results indicate the importance of exploratory methods and distinguishing between approach and avoid components of demonstration goals in the social domain, similar to other domains (Elliot, 2005; Ryan & Shim, 2006).

In sum, the confirmatory factor analysis results provided support for expanding the achievement goal theory into the social domain. Results also indicated that achievement goals are distinct factors across the academic and social domains, and measuring them separately is important.

In terms of convergent validity, the SAGOS was significantly correlated in the expected direction with measures of theoretically related constructs. Specifically, those students that scored high on the SAGOS tended to score high on measures of the AGQ. Correlations among the social achievement goals were similar to those typically found among academic achievement goals. Results of the present validity study give further support to the use of the SAGOS for a variety of educational purposes.

However, despite this initial support for the SAGOS, further studies are needed in several areas. Several additional lines of research are needed to clarify the SAGOS's utility as an assessment tool. The procedure's concurrent validity needs to be investigated further by comparing it to other well-designed instruments that measure a range of achievement goals. Additional studies of the criterion-related validity of the SAGOS are also needed to determine whether the SAGOS can associate with observing student behavior in real-life settings (e.g., quality of interpersonal interactions). Consistent with the findings reported by Elliot and McGregor (2001) in the academic and university settings, the existence of the four-factor social goal model in high school should be studied to thoroughly examine the

mastery-approach, mastery-avoidance, performance-approach, and performance-avoidance factors (Elliot & McGregor, 2001).

Since social mastery, social performance-approach and social performance-avoidance subscales showed a positive relationship, perhaps a line of research is needed to clarify multiple social goals. Students can hold multiple goals at one time (Ames & Archer, 1988; Harackiewicz, Barron, Taauer, Carter, & Elliot, 2000). For example, a high-achieving student could simultaneously be mastery-oriented [this is an interesting problem (or an interesting relationship) that I want to solve (to establish)] and performance-approach-oriented [I want the others to see how good I am at solving the aforementioned problem (establishing the aforementioned relationship)]. Another student could be mastery-oriented and performance-avoid oriented [if I can't solve (establish) this problem (relationship), then I'm going to look stupid]. A competition can encourage students' adoption of performance-approach and performance-avoidance goals, depending on how well they think they will do in the competition. Similarly, in social setting a social competition (e.g. select friends) can encourage students to adopt social performance-approach and social performance-avoidance goals, depending on how well they think they will do in the competition.

The first limitation of the study, as previously noted, was that the sample was comprised solely of high school students. Hence, the findings may be limited in their generalizability to other populations. The second limitation of the current study is its reliance on the use of self-report measures to gather validity evidence, which can create a number of inherent problems, such as measurement error and participant bias. The third limitation was that this study was conducted only in Saveh city. Conceivably, findings may not be generalized to students in other regions of Iran or other countries. The last limitation of the current research was that it was conducted on only one sample; it is recommended that validity results be examined on other samples.

References

- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Student's learning strategies and motivational processes. *Journal of Educational Psychology, 80*, 260-267.
- Anderman, E. M., & Anderman, L. (1999). Social predictors of changes in students' achievement goal orientations. *Contemporary Educational Psychology, 25*, 21-37.
- Covington, M. (2000). Goal theory, motivation, and school achievement: an integrative review. *Annual Review of Psychology, 51*, 171-200.
- Deci, E., & Ryan, R. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 227-268.

- Dowson, M., & McInerney, D. M. (2001). Psychological parameters of students' social and work avoidance goals: A qualitative investigation. *Journal of Educational Psychology, 93*, 35-42.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*(10), 1040-1048.
- Elliot, A. J. (2005). A conceptual history of the achievement goal construct. In A. Elliot & C. Dweck (Eds.), *Handbook of competence and motivation*. New York: Guilford.
- Elliot, A. J., & Dweck, C. S. (2005). Competence as the core of achievement motivation. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation*. New York: Guilford.
- Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*, 501-519.
- Elliot, A. J., & Thrash, T. M. (2001). Achievement goals and the hierarchical model of achievement motivation. *Educational Psychology Review, 13*, 139-156.
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 169-189.
- Elliot, A. J., & Church, M. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*, 218-232.
- Harackiewicz, J., Barron, K., Tauer, J., Carter, S., & Elliot, A. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology, 92*, 316-330.
- Harter, S. (1982). The perceived competence scale for children. *Child Development, 53*, 87-97.
- Hopkins, N., & Ryan, A. M. (2000). Expanding achievement goal theory into the social domain. Poster presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Jöreskog, K., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software International.
- Jöreskog, K., & Sörbom, D. (1981). Analysis of linear structural relationships by maximum likelihood and least squares methods. Research Report 81-8, University of Uppsala, Sweden.
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91*, 328-346.
- Patrick, H., Anderman, L., & Ryan, A. M. (2002). Social motivation and the classroom social environment. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 85-108). Mahwah, NJ: Lawrence Erlbaum.

- Pintrich, P., & Schunk, D. (2002). Chapter five: The role of goals and goal orientation. *Motivation in education: Theory, research and applications* (2nd ed.) Upper Saddle River, NJ: Merrill.
- Rubin, K. H., Coplan, R. J., Nelson, L. J., Cheah, C. S., & Lagace-Seguin, L. J. (1999). Peer relationships in childhood. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (451-502). Mahwah, NJ: Lawrence Erlbaum.
- Ryan, A. M., & Hopkins, N. B. (2003). *Achievement goals in the social domain*. University of Illinois, Urbana-Champaign, unpublished manuscript.
- Ryan, A. M., and Shim, S. S. (2006). Social Achievement Goals: The Nature and Consequences of Different Orientations Toward Social Competence. *Personality and Social Psychology Bulletin*.
- Urdan, T. C. (1997). Examining the relations among early adolescent students' goals and friends' orientation toward effort and achievement in school. *Contemporary Educational Psychology*, 22, 165-191.
- Wentzel, K. R. (2000). What is it that I'm trying to achieve? Classroom goals from a content perspective. *Contemporary Educational Psychology*, 25, 105-115.
- Yu, C., & Muthén, B. (2002). Evaluation of model fit indices for latent variable models with categorical and continuous outcomes. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Simple Formative Assessment, High Learning Gains in College General Chemistry

Yalçın Yalaki*

Suggested Citation:

Yalaki, Y. (2010). Simple formative assessment, high learning gains in college general chemistry. *Egitim Arastirmalari - Eurasian Journal of Educational Research*, 40, 223-241.

Abstract

Problem Statement: The utility of formative assessment for improving student performance is acknowledged in the literature. For formative assessment to be useful, quality feedback, the proper use of feedback by students and revision of practice to accommodate formative assessment is necessary. However, in the contexts of university level science courses, where, in many cases, there is a heavy load of teaching and assessment on both students and instructors, providing quality feedback, using feedback properly, and changing practice may not be as forthcoming as hoped.

Purpose of Study: The purpose of this study is to show that providing opportunities for formative assessment, even at a relatively simple and manageable level, could significantly impact students' achievement and also their attitude toward university level science classes without a considerable change in practice.

Methods: This study was conducted in an undergraduate general chemistry course with an emphasis on organic chemistry taken by pre-service middle school science teachers over the period of two spring semesters in two consecutive years. In total, 163 students participated in the study. A qualitative research methodology accompanied quantitative methods for more in-depth understanding. Summative exam results, responses to a questionnaire, observations, and interview transcripts provided the data for the study.

Findings and Results: Quantitative results showed statistically significant improvement in experimental group students' exam grades, which was an indication of improvement in achievement and learning. Through qualitative data, positive student reactions toward the formative

* PhD, Hacettepe University, Faculty of Education, yyalaki@hacettepe.edu.tr

assessment process used in the study were observed. Insight into students' perceptions of the formative assessment methods and their use of feedback were obtained.

Conclusions and Recommendations: The relatively simple formative assessment process utilized in this study has the potential to improve both student achievement and learning and also students' attitudes toward the courses they take. Therefore, university level instructors, who teach science courses, should consider using formative assessment methods, at least at a simple level without a too heavy a load of extra work, to improve student achievement and attitude. However, it is not possible to claim that formative assessment methods will improve student achievement and attitude in all cases.

Keywords: Formative assessment, formative feedback, peer-feedback, self-assessment, science education, higher education

The utility of formative assessment for improving student learning is acknowledged in the literature. Widely cited publications about formative assessment such as the meta analysis by Black and Wiliam (1998) and a book published by Organisation for Economic Co-operation and Development [OECD] (2005): *Formative assessment: improving learning in secondary classrooms*, claim that formative assessment has a high potential of providing positive learning gains among other educational innovations. However, applying formative assessment in classrooms is a challenge for many teachers and instructors, especially in higher education (Carless, 2007; Yorke, 2003). The purpose of this study is to demonstrate that in classrooms without severe motivation challenges, even simple applications of formative assessment can be highly successful in terms of student attitude and learning.

According to the CERI (2005) book "Formative assessment refers to frequent, interactive assessments of student progress and understanding to identify learning needs and adjust teaching appropriately (p. 21)." Successful applications of formative assessment have several features. Nicol and Macfarlane-Dick (2006) identify the following principles of good feedback practice in formative assessment:

Good feedback practice:

1. helps clarify what good performance is (goals, criteria, expected standards);
2. facilitates the development of self-assessment (reflection) in learning;
3. delivers high quality information to students about their learning;
4. encourages teacher and peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape teaching (p. 205)

These principles summarize what features an ideal feedback practice in formative assessment would have. However, there are many factors that affect the success of

formative assessment such as learning context, student motivation, student-teacher relations, and students' level of cognitive development (Elwood 2006; Higgins, Hartley & Skelton, 2002; Perrenoud 1998; Yorke 2003). These factors influence how many of the principles above could be utilized effectively by instructors and students. Higgins et al. (2002) ask important questions about student motivation and the practice of feedback:

For formative assessment to work in practice, feedback must 'connect' with students. But, at a time when student numbers are rising and competition for graduate jobs is growing, are students increasingly becoming instrumental consumers, driven by the extrinsic motivation of the mark? If so, will they heed written feedback which encourages them to reflect on their learning? Or will they simply pay attention to the grade, and seek feedback only when it is perceived to provide 'correct answers' to commit to memory (and only then when their grade expectation has not been met)? (p. 54)

They answer these questions by arguing that even if grades and competition for jobs are a source of motivation for students, students do use feedback in many different ways. Perrenoud (1998) expresses the importance of learners in formative assessment, reminding us that the idea of feedback and formative assessment would be incomplete without the learner and that 'feedback does not in itself guarantee better learning.

Feedback is a simple message. How can it assist the learning process? Because pupils take it into account, because it affects their cognition. Communication theory teaches us that the effectiveness of a message is measured at the level of the recipient: an intervention or a piece of information only helps a pupil learn better if their thought processes are modified. This is an abstract way of stating that no learning takes place without the learner. One can only stimulate, reinforce, reorient, readjust or accelerate the pupil's mental processes, in the hope of modifying the learning processes. (p. 86)

Yorke (2003) agrees by arguing that "the importance of the student's reception of feedback cannot be overstated." (p. 488) Students' goals, motivation, personality orientation, self-efficacy and other factors influence how they receive and use the provided feedback. For example, self-efficacy plays an important role on how individuals regulate the use of feedback, among other things, which varies from person to person (Akbaş, 2010; Erdem 2008). On the other hand, providing quality feedback is a challenge for teachers and instructors. As Black and Wiliam (2009) argue, teachers usually have only a few seconds to interpret student contributions and choose an optimum response. From these arguments, it is possible to conclude that using formative feedback is a complex process on the side of both students and teachers and does not guarantee improved learning. As Yorke (2003) argues, establishing a direct connection between formative assessment and learning and achievement warrants caution. Black, Harrison, Hogden, Marshall and Wiliam (2005) also agree that formative assessment does not guarantee improvement in student achievement in all cases.

We do not claim that formative assessment leads to improved student achievement in *all* cases, with *all* teachers on *all* occasions. ... Our claim is that formative assessment in *general* is an effective intervention, although we do not estimate the difficulties in translating theory into practice. (p. 7, original emphasis)

Carless (2007) argues that "In particular, large class sizes and heavy workloads often present a barrier to teachers' implementation of formative assessment." (p.

173). He recognizes that "...there is a need to find ways to make formative assessment more attractive and manageable for practitioners..." (p. 174). Higgins et al. (2002) and Yorke (2003) also mention the problems regarding the implementation of formative assessment. They also argue that in higher education, the challenges that instructors face in utilizing formative assessment include high student/instructor ratio, other instructor responsibilities than teaching, and emphasis on summative assessment by the institutions. Hence, practicing formative assessment by providing quality feedback, using feedback on the part of students, and adjusting teaching may not be a simple matter in higher education.

Recognizing the difficulties faced in the practice of formative assessment, many authors still agree that formative assessment is a part of good teaching practice (Ahmed & Teviotdale, 2007; Black et al., 2005; Black & Wiliam, 2009; Carless, 2007; Sadler, 1998; Yorke, 2003). This study is based on the argument that formative assessment has a potential to improve student learning. The problems faced in utilizing formative assessment could be surmounted by simple approaches. In order to keep instructor workloads relatively unchanged while implementing formative assessment practices in teaching, some choices could be made on the type of feedback to utilize. The feedback in formative assessment can take different forms such as, written feedback, oral feedback, grading, peer feedback and self-assessment (Sadler, 1998; Black, 2003; Nicol & Macfarlane-Dick, 2006). All these feedback forms have their advantages and disadvantages. For example, written feedback may be more time consuming for instructors, while oral feedback can be given more easily and in a more timely manner. Peer feedback and self-assessment provide better reflection on the part of learners (Liu & Carless, 2006). Encouraging peer dialogue during the formative assessment process promotes a sense of ownership among students while enhancing learning (Nicol, 2007). In this study, oral feedback, peer feedback and self-assessment were utilized, because they required less work on the part of instructors. This decision was important since the purpose of this study was to demonstrate that least amount of additional work on the part of instructors, who are already busy with other occupations, could still stimulate useful feedback practices. Besides, an important role of formative assessment is to improve student self-regulation in learning, which would be improved by peer feedback and self-assessment (Nicol, 2007 and 2009).

The research questions of the study were:

- 1- Could formative assessment improve pre-service middle school teachers' achievement in chemistry classes?
- 2- What are the perceptions of pre-service middle school teachers toward the utilization of formative assessment in chemistry classes?
- 3- How do pre-service middle school teachers use feedback in a formative assessment process?

A mixed qualitative and quantitative research methodology was used in this study with a strategy of case oriented action research. Even though these two methodologies have very different assumptions about the phenomena to be investigated, mixed methodologies involving both quantitative and qualitative research are advocated and used in educational research (Johnson & Onwuegbuzie,

2004). The reason for choosing this mixed methodology for this research was the richness of understanding achieved by the complementary use of the two methods at the same time. Sale, Lohfeld and Brazil (2002) argue that these two methodologies, if used together, should be used in a way that they complement each other.

In this study, the phenomena investigated by quantitative and qualitative approaches can be differentiated easily. The quantitative data provided information on the general progress and views of participants while qualitative data provided more in-depth and enhanced understanding of their perceptions of the formative assessment used in this study and how they utilized it. Based on Johnson and Onwuegbuzie's (2004, p.22) conceptualization of mixed methodology, this study used quantitative and qualitative methods concurrently and with equal emphasis.

Method

The study was conducted at a major university in Turkey in a general chemistry course, with an emphasis on organic chemistry, taken by second-year pre-service middle school science teachers over the period of two spring semesters in two consecutive years. 163 students participated in the study. The study was designed in a way to compare the achievement scores of students who took the class in 2007 with those who took it in 2008. The 2007 class received regular instruction without a formative assessment process and limited feedback while the 2008 class received instruction with a formative assessment process and deliberate feedback. The 2007 class consisted of two groups with 47 and 48 students in each group. This grouping was done to better manage classroom space and class schedules. Both groups received two midterm exams and one final exam. The 2008 class also consisted of two groups for the same reasons each containing 34 students. Instead of conducting the research with two groups in the same year, the reason for this sequential design was to prevent control and treatment groups from interacting with each other and also preventing a feeling of unfairness among students. The students were assigned to the groups based on the alphabetical order of their last name, which provided a level of randomization in group assignment of individuals. However, convenience sampling rather than random sampling was employed in this study, since the purpose of the study was reporting a case and not generalizing the results to a population.

The 2008 class received four quizzes without grading as treatment for the purpose of providing formative feedback to students on top of the required two midterm exams and one final exam. Feedbacks on the midterm exams were also provided to the 2008 class. Midterms and final exams were used as a measure of achievement for both classes. The tests applied for the midterms and final-exams provided the quantitative data for the study. Reliability and validity studies were completed for all of the tests applied. Table 1 shows the reliability statistics of the tests. As seen in the table, because of conflicting schedules, the groups in the 2007 class received different tests, but they were similar in structure and content, with same number of items. The groups in the 2008 class received the same midterm and final tests. 2007 and 2008 tests contained different but similar items and the format and structure of all the tests were the same. The test items were classical paper and pencil items, most of which required short answers. However, in order to calculate a

reliability score, all of the test items were coded as “right” or “wrong” even though partial credits were given for partially correct answers. The point worth of items varied, but, for example, for a test item worth 10 points, if a student got 6 points or more for the item, the item was coded as “right,” and if the student got less than 6 points, then the item was coded as “wrong.” Even though the results averaged out to some degree, some information was lost in this process regarding the grades of the students. However, all of the tests applied provided high reliability scores at the end of this process, as shown in Table 1.

Table 1
Reliability Statistics of the Applied Tests

	2007 Class				2008 Class	
	Group A (47 Std.)		Group B (48 Std.)		Group A (34 Std.)	Group B (34 Std.)
	Number of questions	KR21	Number of questions	KR21	Number of questions	KR21
Midterm 1	24	0.80	24	0.82	25	0.74
Midterm 2	17	0.84	17	0.82	21	0.79
Final	23	0.79	23	0.73	30	0.71

In order to achieve content validity for the tests, a table of the course objectives was developed. The number of questions that needed to be asked for each objective, and at what level, was decided prior to the administration of the tests. Most of the questions in the test measured procedural knowledge, which required application of some rule, while other questions measured declarative knowledge (Oosterhof, 2001).

Another instrument that was developed for this study was a questionnaire with a Likert type scale to measure students' interests toward organic chemistry and the class they took, as well as their reactions toward the formative assessment process utilized, which contained 26 items plus four open ended questions. Each item had five possible responses ranging from “Do not agree at all” to “Totally agree.” The students were asked to write a pseudonym rather than their names on the questionnaire papers to ensure anonymity. A frequency analysis of the items in the instrument was performed after gathering students' responses in a pilot study; the opinions of four colleagues about the items were solicited, and some revisions were made based on these findings to ensure validity. The reliability score of the instrument was found to be Cronbach alpha = 0.896. At the end of the study, 66 students from the 2008 class had returned the instrument with their responses. The questionnaire provided both quantitative and qualitative data regarding students' views, reactions and perceptions.

As the author and investigator of this study, I taught all of the courses, conducted the quizzes and exams and also evaluated them. As mentioned before, a case-oriented action research strategy was employed to conduct the study. I was a participant observer and investigator at the same time. During the formative assessment process, I took notes about my observations related to students' use of the formative feedback that they got from their peers and me as the instructor.

Finally, semi-structured interviews were conducted with seven volunteer students to collect further data on students' perceptions of the formative assessment process. Both the open ended questions in the questionnaire and interviews provided detailed data on the students' opinions about the formative process utilized in the class and also how they used the formative feedback generated during this process. The interviews were recorded and transcribed for data analysis.

Data Analysis

As mentioned above, quantitative and qualitative data were collected in this study. The administered test results and the scores obtained from a Likert type questionnaire served as the quantitative data. For analyzing these data, an independent sample t-test and a frequency analysis were used, respectively. Answers given to the open ended questions in the questionnaire, observation notes, and interview transcripts made up the qualitative data. To analyze the qualitative data, codes were developed in order to find themes and patterns in the data (Bogdan & Biklen, 2002), which helped with interpreting the data and also with the process of making meaning out of data.

Findings and Discussion

The quantitative data gathered through the applied tests revealed a significant difference in the scores of both classes. When the test results of the combined groups for each class were compared, it was found that the 2008 class, in which formative assessment was utilized, had significantly higher scores in all of the tests. Table 2 show the related statistics with the assumption that the homogeneity of variance is not violated for all the cases.

Table 2

Group statistics (Chemistry 4 test results are out of 100)

	Class	N	Mean	Std. Dev.	Std. Error Mean	p (2-tailed sig.)
Midterm 1	2008	68	87.3	10.38	1.26	<.001
	2007	95	66.0	19.18	1.97	
Midterm 2	2008	68	76.5	18.16	2.20	<.001
	2007	95	63.9	23.66	2.43	
Final	2008	68	68.8	15.53	1.88	.010
	2007	95	61.4	19.45	2.00	

No pre-test was administered in this study, because there was evidence that showed all of the participating students had similar competencies in the beginning. When their average achievement scores for the previous three chemistry classes that they took prior to this study are compared, as shown in Table 3, there is no significant difference. The first two previous chemistry courses were taught by the same instructor for both groups, with similar tests in content and design. The third previous chemistry course was taught by the author of this study to both groups, again with similar tests in content and design. The tests applied to both groups in these three classes were similar enough in content and design to allow comparison, but the questions had to be different enough to prevent 2008 class students from having an advantage over the 2007 class students by collecting information on previous tests. This information was considered enough to assume that the compared groups had similar achievements in chemistry in the beginning.

Table 3

Comparison of average achievement scores of the previous three chemistry classes

2007 Class	2008 Class
average of the achievement scores for Chemistry 1, 2, and 3 classes	average of the achievement scores for Chemistry 1, 2, and 3 classes
73 (out of 100)	74 (out of 100)

When it comes to the questionnaire data, the 26 items in the questionnaire were designed to collect information about students' perceptions about the course that they were taking, the content of the course, the instructor, and also the formative assessment practices used in the course. As noted earlier, 66 participants returned the questionnaire with their responses. The frequency analysis of the responses is shown in Table 4. The responses are given either as the total of positive responses (agree + strongly agree) with a + mark or as the total of negative responses (disagree + strongly disagree) with a - mark.

Table 4*Student responses to the items on the questionnaire (n=66)*

	Questionnaire Items	Responses (%)
1	This course will be useful to me in my job in the future.	+80.3
2	This course includes information that I can use in daily life.	+75.8
3	I value what I learn in this course.	+90.9
4	If we didn't take this course, we would not loose anything.	-86.4
5	The taught subjects were in agreement with the course objectives	+94.0
6	The way this course was taught was in agreement with the course objectives	+81.8
7	I did not fully learn the taught subjects	-59.1
8	I am interested in Organic Chemistry.	+63.6
9	The instructor was not very competent in course subjects	-92.4
10	Organic chemistry is over my capacity of comprehension	-84.8
11	This course could have been taught in a better and clearer way	-57.6
12	This course has no contribution to my job training	-80.3
13	The course content was not suitable for my expectations	-86.4
14	I find organic chemistry totally uninteresting	-80.3
15	The instructor gave the course in a way that was clear and easy to understand	+93.9
16	I did not learn anything useful in this course	-80.3
17	I learned the subjects well that was discussed in the course	+77.3
18	The instructor was insufficient in terms of teaching techniques	-77.3
19	The instructor knows organic chemistry very well	+92.4
20	Organic chemistry is one of the easy classes for me	+40.9
21	The quizzes applied in this course were very useful.	+94.0
22	The quizzes applied in the course helped me learn.	+95.5
23	There should have been more quizzes.	+42.4
24	Quizzes made no contribution to my learning.	-97.0
25	Nothing would have changed if no quizzes were applied.	-97.0
26	Having too many quizzes was boring.	-92.4

The findings in Tables 2 and 3 imply a higher level of chemistry achievement in the 2008 class. However, it is not possible to claim that the sole reason for the improvement observed in the achievement scores of the 2008 class is the formative assessment process. There are many variables that influence a group of students' achievement that are not possible to control. For example, one factor that may have affected the achievement of students is the reduced class sizes in the 2008 group. Other factors may include the frequent changes in the central exam and student placement system for university entrance in Turkey. The content of questions asked in the national university entrance examinations, the rules of these examinations, and the number of students accepted to a particular program each year have changed over the years, which may have influenced the backgrounds of students who entered university in different years. However, the percentile range of students that enter the particular program in the university where this study was conducted did not change significantly over the years. This provides some confidence about the conclusion that at least some of the increased achievement scores of the 2008 class, when compared to the 2007 class, come from the utilized formative assessment process.

Students' very positive reactions toward the use of quizzes for formative purposes support this conclusion to some extent. As seen in Table 4, the vast majority of the students thought that the applied quizzes contributed to their learning. They also had mostly positive attitudes toward the class and its content, which is a crucial issue that influences how much they value and utilize the formative assessment process. Without this positive attitude it would be much more difficult for an intervention to have a positive affect. The reason why students had positive attitudes toward the course is because they all have a relatively strong science background, since they chose to study in the scientific field in high school. The positive reactions of students toward the formative assessment process that was used in this study were also reflected in their responses to some of the open ended questions in the questionnaire that was given to them.

As indicated before, quizzes were given to students for formative purposes. Before giving examples of students' responses, the application of quizzes should be made clearer at this point. As explained in the methods section, four quizzes were administered in the 2008 General Chemistry class during one semester. The first two quizzes were administered before the first midterm examination, while the last two were administered before the second midterm examination. Students were told that the quizzes were not going to be graded in order to focus their attention on the learning aspect of the quizzes rather than the grades. Not grading the quizzes also helped with decreasing the grade-related stress on the part of students. Also, not having to grade the quizzes kept the workload of the instructor virtually unchanged. The quizzes were given to students for a period of about 30 minutes and during this time they were encouraged to solve the quiz questions on their own to see where their weaknesses and misunderstandings were. After this time, students were asked to swap papers with a friend and compare their answers. Then those who swapped papers were encouraged to discuss with each other and try to find out the reasons for their different answers, if any. During this time, students also discussed the answers in small groups and had an opportunity to interact with their friends' ideas about quiz questions. They were able to hear their friends' answers for questions that they could not answer. After this period, a break was given, and when the students came

back, the instructor answered questions during a whole class discussion. During the small group and class discussions, the instructor encouraged the students to find their weaknesses and misunderstandings. In order to aid students in improving their learning, the instructor adjusted the teaching by making further explanations, giving different examples, and solving more problems instead of just continuing onto the next subject.

The open-ended questions related to the given quizzes in the questionnaire were: "What are your thoughts about the quizzes given in the class? Do you think that quizzes contributed to your learning? How should the quizzes have been done to be more useful?" Almost all students gave positive answers to these questions. One example of a student response was:

I think that quizzes given in the class were definitely useful, because they allowed us to see where we are lacking and our mistakes before the course advanced further. This way we could complete our knowledge in these subjects. Besides, thanks to the quizzes we were able to solve different examples. The way quizzes were conducted was nice. (Student response to open ended questions)

Another students' response was:

Students usually study from exam to exam and because of this many subjects remained uncomprehended before the exams. Thanks to quizzes, these subjects were better understood. (Student response to open ended questions)

There were many more examples similar to the quotes above. Besides quizzes, similar feedback was given about the midterm examinations as well. Students had a chance to review their exam papers, compare their answers with their friends and finally learn from their mistakes during whole class discussions. In this process, students performed self-assessment and were also exposed to peer feedback and instructor feedback. As a result of this process, some subjects, where many students had shortcomings, were taught and explained in different ways with different examples. This was also a unique experience for the instructor, who did not have a chance to see and intervene with his students understanding and learning to this extent before.

Interviews provided further insight into students' perceptions of the formative process used in this study and also how they made use of the feedback that they got from peers and the instructor. One of the findings that came out of the interview transcripts had to do with the influence of the quizzes that was used for formative assessment on how often the students studied for the class. When asked "How often do you study for this class at home?" one of the students, Ferda (pseudonym), explained:

As often as the quizzes... Quizzes were very good, in that they made this class one of the easiest in terms of preparing for the midterms and the final exam. Because we did not wait till the last moment to study. You know, in general it is student psychology to study for an exam only a week before, but this class wasn't like that. (Interview transcripts with Ferda 22.05.2008)

Another student, Tulin's (pseudonym) response to the same question was:

I was studying right before the quizzes. Other than that I was taking regular notes in the class, I have all the notes. ...[The quizzes] definitely influenced. ... Since no grade was

given, nobody memorized as they do before the exams. We studied for learning really. (Interview with Tulin, 21.05.2008)

Finally, Yelda (pseudonym) had this to say about this question:

... I will generalize about all students; we leave everything to the last day. If there were no quizzes, we probably would have studied chemistry only the last day [before the exams] but since there were quizzes, we at least looked at what we have done in the quizzes, we could see where we were lacking, and we could specify which subjects to study. This, I believe, encouraged us to study gradually; we would not just go to the exam and feel disappointed, we knew what we know and what we don't. (Interview with Yelda, 26.05.2008)

These responses show that quizzes did effect how often students studied for the class. Most students usually study for a class just before the exams, but the quizzes applied in this study appeared to have increased the frequency of study and review on the part of students. Another issue that the interview findings revealed was that most students attach great importance to peer-review during the application of quizzes. One of the students, Mine (pseudonym), said,

I think reviewing the quizzes among ourselves is very useful, because, it matches human psychology, what your friend says may be more effective and longer lasting in memory than what the teacher says, because our way of thinking is similar. I think it was necessary, I mean, it was logical to review with our peers. (Interview with Mine, 22.05.2009)

Another student, Aynur (pseudonym) agreed,

I think sharing is better... It was good to discuss with friends. I think I learned better. The more I hear information from my friends, the better. No matter how much I study on my own, if I can't learn something, I can't learn it. But when there is interaction, everything is faster. (Interview transcripts with Aynur 21.05.2008)

Tulin also agreed and she argued that it is better to do peer-review than self-assessment, because this way, students cared more about what they did in the quizzes.

It was good to do exchanges among friends. ... I learned there. ... When we gave [our papers] to friends, it was more effective, because, I don't know, maybe it is this worry that my friend will think that I don't know anything, but at that time there was this psychology of trying to do more, since somebody else is going to review it; that was very effective. It is good to have friends to evaluate you, otherwise when there is only self-assessment, my thinking is "I am going to study anyway" and people may care less. (Interview with Tulin, 21.05.2008)

One student, Uhde (pseudonym), however, did have reservations about peer-review. She said,

...There were things that I learned incorrectly during the review of the first quiz. We discussed with friends, I placed it in my memory as I learned this, as if it was right. As I was studying, I looked at the exam and said, "This is how I learned it." I mean, instead of discussing with friends, well, maybe we should do that too, but it was more effective for me when you solved the questions on the board. (Interview with Uhde, 27.05.2008)

Finally, another student, Umit (pseudonym) said this about peer-review,

If we knew the subject, we could see where our friends were lacking, if we didn't, we could say "this is how my friends did this, why didn't I think about that. (Interview with Umit, 26.05.2008)

From these findings, it is obvious that most students in this study valued peer-review and they saw this process as something that aided their learning. All of the interviewed students, except Mine, thought that it was good that no grade was given to the quizzes. They said that not giving quizzes took away the stress associated with grades and expressed the learning aspect of the quizzes. Mine, on the other hand, thought that giving at least partial credit to the quizzes could have made it more important for students and they would have taken it more seriously. Others disagreed. Mine's comments about giving grades on the quizzes came after the question related to the last two quizzes. During the last two quizzes, students seemed less interested and less engaged. When I asked the reason, Mine thought that if grades were given to the quizzes, students would have taken them more seriously. However, all other students thought that not giving grades was not the problem. The problem was that during the last two quizzes students were so busy with other homework and exams that they were just exhausted. They thought that giving grades to the quizzes would have made it a negative experience; students would have stressed out on top of being exhausted and this would not have helped their learning at all.

Another issue that was raised during the interviews was the importance of quizzes. When students were asked, "What is the most important contribution of the quizzes, helping you learn or helping you to improve your grades?" all of the students responded that both are important. Therefore, at the college level, where grades are very important for students for many good reasons, it is possible to conclude that learning is not easily detached from grades. This poses a challenge to formative assessment, since the purpose of formative assessment is to improve learning rather than giving summative grades. This is why no grades were offered during the formative assessment process during this study.

When it comes to the types of questions that were asked in the quizzes, students had mixed responses. While most agreed that the quiz questions were fine, there were suggestions to increase the types of questions or add questions that were more related to daily life. Five of the interviewed students thought that quiz questions should be similar to the exam questions, while two of them said that this was not necessary. For example, Ferda thought that if the quiz questions were very different than those on the exam, the value of the quizzes would have been diminished. This thinking is in line with the importance attached to exam grades and exam preparations. Most students thought that exam preparation was extremely important in addition learning.

Observation notes revealed that most students were engaged during the quizzes, discussing with each other while answering questions. Observation and interview data also revealed that students used feedback that came from their peers or the instructor to review their learning. There were always a few students who were disengaged during the quizzes for various reasons. The number of disengaged students increased as the semester progressed, and as the interviews revealed, this was due to the higher work load for students.

Conclusion

It can be argued from the evidence presented in this study that formative assessment, even at a relatively simple level, could significantly improve achievement in science classes in higher education, if instructors are prepared to utilize it. Even though there may be other factors that influenced the improved achievement scores in this study, the formative assessment process seemed to have a positive contribution. Many studies confirm this positive contribution in higher education in many different fields such as health sciences (Carrillo-de-la-Penã et al., 2007), medicine (Kibble, 2007; Dobson, 2008), biology (Peat and Franklin, 2002), psychology (Costa, Mullan, Cothe & Butow, 2010), and math (Lawson, 1999). Kibble (2007) and Dobson (2008) report using online quizzes as formative assessment tools with positive results. Most of the studies above used online or computerized tests as tools for formative assessment and the point of doing so was to reduce the workload on the instructor while utilizing formative assessment.

The formative assessment process used in this study was also designed in a way to prevent adding too much workload onto the busy instructors while still providing further learning opportunities for students. From the findings of this study, it could be asserted that a formative assessment process using quizzes or short exams as a means to assess student learning and provide feedback should take the following factors into consideration:

- No grades, even partial credit, should be given to quizzes. This helps focus students' motivation on the learning aspect rather than grades.
- Peer feedback is important and provides students opportunities to interact with their peers while learning. Sometimes peers use language that is easier to understand for other students, and may aid learning.
- Quizzes increase the frequency of assessment and the frequency of study for students, which seem to benefit learning.
- Grades on exams are as important as learning, if not more, for students; therefore, one of the incentives of quizzes for students may be that they help improve grades. This could be used to the advantage of the formative assessment process, in that students could attach more value to the quizzes instead of seeing them as mere exercises.

This study is an example of a successful application of formative assessment without significant changes to teaching practices. However, it is not possible to claim that this process would be successful in any college level science class. One of the important issues that need to be considered is student motivation and self-efficacy. Without student motivation or value placed by students on the process, it is very difficult for formative assessment to make a difference. For further research, I suggest exploring other simple methods of utilizing formative assessment in undergraduate and post-graduate levels that do not place too great an extra workload on the instructors. Otherwise, formative assessment may not find widespread application at the college level.

Acknowledgments

I would like to thank Prof. Fitnat Kaptan for her valuable help and suggestions during the study and the data analysis process. I also would like to thank Dr. Ali Ekber Sahin, Dr. Serkan Yilmaz, Dr. Pinar Ozdemir, and Bilge Gök for their help during the instrument development and data analysis.

References

- Ahmed, N. & W. Teviotdale (2007). The Value of Formative Assessment in Higher Education. Education in a Changing Environment, Greater Manchester, UK, Informing Science Press.
- Akbaş, A. (2010). Attitudes, self-efficacy and science processing skills of teachign certificate master's program (ofmae) students. *Egitim Arastirmalari - Eurasian Journal of Educational Research*, 39, 1-12.
- Black, P. (2003, April). *Formative and Summative Assessment : Can They Serve Learning Together?* Paper presented at the Amercian Educational Research Association, Chicago, IL.
- Black, P., Harrison, C., Hogden, J., Marshall, B., & Wiliam, D. (2005). Dissemination and evaluation: a response to Smith and Gorard. *Research Intelligence*, 93, 7.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-73.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Bogdan, R. C., & Biklen, S. K. (2002). *Qualitative research in education*. Massachusetts: Allyn and Bacon.
- Carless, D. (2007). Conceptualizing pre-emptive formative assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), 171-184.
- Carrillo-de-la-Pena, M. T., E. Bailles, Caseras, X., Martinez, A., Ortet G. & Perez, J. (2009). "Formative assessment and academic achievement in pre-graduate students of health sciences." *Advances in Health Sciences Education* 14(1): 61-67
- CERI. (2005). *Formative assessment: Improving learning in secondary classrooms*. Paris: OECD.
- Costa, D. S. J., B. A. Mullan, et al. (2010). "A web-based formative assessment tool for Masters students: A pilot study." *Computers & Education* 54(4): 1248-1253.
- Dobson, J. L. (2008). "The use of formative online quizzes to enhance class preparation and scores on summative exams." *Advances in Physiology Education* 32(4): 297-302.

- Elwood, J. (2006). Formative assessment: possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy & Practice*, 13(2), 215-232.
- Erdem, M. (2008). The effects of the blended teaching practice on prospective teachers' teaching self-efficacy and epistemological beliefs. *Eurasian Journal of Educational Research*, 30, 81-98.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The Conscientious Consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53-64.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(14), 14-26.
- Johnson, R. B. and A. J. Onwuegbuzie (2004). "Mixed Methods Research: A Research Paradigm Whose Time Has Come." *Educational Researcher* 33(14): 14-26.
- Kibble, J. (2007). "Use of unsupervised online quizzes as formative assessment in a medical physiology course: effects of incentives on student participation and performance." *Advances in Physiology Education* 31(3): 253-260.
- Lawson, D. (1999). "Formative assessment using computer-aided assessment." *Teaching Mathematics and Its Applications* 18(4): 155-158.
- Liu, N., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290.
- Nicol, D. J. (2007). Principles of good assessment and feedback: Theory and practice. International Online Conference sponsored by the REAP Project: Assessment design for learner responsibility. Retrieved April 05, 2010, from http://www.reap.ac.uk/public/Papers/Principles_of_good_assessment_and_feedback.pdf
- Nicol, D. J. (2009). "Assessment for learner selfregulation: Enhancing achievement in the first year using learning technologies." *Assessment & Evaluation in Higher Education* 34(3): 335-352.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and selfregulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Oosterhof, A. (2001). Classroom applications of educational measurement. Upper Saddle River, N.J., Merrill.
- Peat, M. and S. Franklin (2002). "Supporting student learning: The use of computer-based formative assessment modules." *British Journal of Educational Technology* 33(5): 515-523.

- Perrenoud, P. (1998). From Formative Evaluation to a Controlled Regulation of Learning Processes. Towards a wider conceptual field. *Assessment in Education: Principles, Policy & Practice*, 5(1), 85-102.
- Sadler, D. R. (1998). Formative Assessment: revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77-84.
- Sale, J. E. M., Lohfeld, L. H., & Brazil, K. (2002). Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research. *Quality & Quantity*, 36, 43-53.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45, 477-501.

Üniversite Genel Kimya Derslerinde Basit Biçimlendirici Değerlendirme, Yüksek Öğrenme Kazanımları

(Özet)

Problem Durumu: Literatürde biçimlendirici değerlendirmenin öğrencilerin performansını artırmadaki yararı çokça tartışılmış ve genel olarak kabul görmüştür. Biçimlendirici değerlendirme öğrenci gelişiminin ve algısının sıkça, etkileşimli olarak değerlendirilmesi ve öğrencilerin ihtiyaçlarının belirlenerek öğretimin buna göre yeniden düzenlenmesi olarak tanımlanabilir. Biçimlendirici değerlendirmenin yararlı olabilmesi için kaliteli dönütler verilmesi, bu dönütlerin öğrenciler tarafından uygun bir şekilde kullanılması gereklidir. Fakat üniversite seviyesindeki fen derslerinde, çoğu zaman, hem öğretmenler hem de öğrenciler üzerinde eğitim-öğretim ve ölçme-değerlendirme açısından yoğun bir yük vardır. Bu nedenle kaliteli dönütler vermek, bu dönütleri uygun bir şekilde kullanmak ve öğretim uygulamalarını değiştirmek beklendiğinden daha zor olabilir.

Araştırmanın Amacı: Bu çalışmanın amacı biçimlendirici değerlendirme için kolay uygulanabilir bir düzeyde fırsatlar oluşturulmasının, öğrencilerin başarılarını ve üniversite seviyesindeki fen derslerine yaklaşımlarını, öğretim uygulamalarında fazla bir değişikliğe gerek kalmadan, olumlu bir şekilde etkileyebileceğini göstermektir.

Araştırmanın Yöntemi: Bu araştırma ilköğretim fen bilgisi öğretmen adaylarının aldığı lisans düzeyindeki organik kimya ağırlıklı bir genel kimya dersi kapsamında yapılmıştır. Dersi 2007 ve 2008 yıllarında alan toplam 163 kişi bu çalışmaya katılmıştır. Araştırmada hem nicel hem de nitel metotlar bir arada kullanılmıştır. Çalışmada 2007 ve 2008 yıllarında dersi alan öğrencilerin başarıları karşılaştırılmıştır. 2007 yılında dersi alan öğrencilere normal öğretim yapılırken, 2008 yılında dersi alan öğrencilere öğrenmeyi desteklemek amacıyla biçimlendirici değerlendirme uygulanmıştır. Biçimlendirici değerlendirme, dönem içerisinde not verilmeden yapılan kısa sınavlar çerçevesinde öğrencilerin kendilerini değerlendirmeleri, akranlarından ve öğretim

elemanından dönüt almaları ve bu dönüt ve değerlendirmeler ışığında öğretimde düzenlemeler yapılması şeklinde gerçekleştirilmiştir. Çeşitli veriler ışığında kontrol ve deney grubu öğrencilerinin başlangıçtaki başarı düzeylerinin aynı olduğu görülmüştür. Öğrencilerin başarılarını ölçmek için güvenilirlik ve geçerlik çalışmaları yapılmış ara sınavlar ve final sınavı kullanılmıştır. Ayrıca geliştirilen Likert tipi bir anket yardımıyla uygulama grubundaki öğrencilerin derse, dersin işlenişine ve ders içeriğine karşı tutumları ve biçimlendirici değerlendirme uygulamasına karşı tepkileri ölçülmüştür. Anketin içerdiği dört adet açık uçlu soruya verilen cevaplar, yedi öğrenciyle yapılan yarı yapılandırılmış mülakatlar ve gözlem notları çalışmanın nitel verilerini oluşturmuştur.

Araştırmanın bulguları: Nicel verilerin analizi uygulama grubu öğrencilerinin başarısında istatistiksel olarak anlamlı bir artışın olduğunu göstermiştir. Her ne kadar bu artışa etki eden ve kontrol edilemeyen başka faktörlerin olması olası ise de, bu artışta kullanılan biçimlendirici değerlendirme uygulamasının da bir katkısı olduğu değerlendirilmiştir. Anket sorularının frekans analizi öğrencilerin derse, dersin içeriğine, dersin işlenişine ve kısa sınavların kullanımına karşı oldukça pozitif tutumlar sergiledikleri görülmüştür. Bu pozitif tutumların biçimlendirici değerlendirme uygulamasının başarısına olumlu bir katkı getirdiği söylenebilir. Anketlerdeki açık uçlu sorulara verilen cevaplar ve yedi öğrenci ile yapılan mülakat çözümlenmeleri öğrencilerin uygulanan biçimlendirici değerlendirme yaklaşımına karşı olumlu tutumlarını teyit etmiş ve öğrencilerin bu süreci nasıl algıladıkları ve kullandıkları ile ilgili bilgiler vermiştir. Bu veriler ışığında öğrenciler genel olarak biçimlendirici değerlendirme amacıyla yapılan kısa sınavlara not verilmemesinin olumlu bir uygulama olduğu, kısa sınavlar sırasında arkadaşlarından dönüt almalarının önemli olduğu, kısa sınavların hem öğrenmelerine hem de notlarına olumlu katkı getirdiği ve kısa sınavların derse çalışma sıklıklarını artırdığı yönünde görüş bildirmişlerdir.

Araştırmanın Sonucu ve Önerileri: Lisans düzeyinde bağıl olarak basit ve uygulanması kolay bir şekilde tasarlanmış olan bir biçimlendirici değerlendirme yöntemi, uygun şartlarda, öğrencilerin başarısını ve aldıkları derslere karşı tutumlarını olumlu bir şekilde etkilemektedir. Bu nedenle lisans düzeyinde eğitim veren öğretim elemanlarının, kendilerine fazla bir yük getirmeden, biçimlendirici değerlendirme yaklaşımlarını kullanmayı düşünceleri önerilmektedir. Biçimlendirici değerlendirme çok farklı şekillerde uygulanabilir, fakat şu hususlara dikkat edilmesi önerilebilir:

- Biçimlendirici değerlendirmede not kullanılmamalıdır.
- Akran dönütlerine ve öz değerlendirmeye fırsat verilmelidir.
- Kısa sınavlar öğrencilerin çalışma sıklığını ve konulara hâkimiyetlerini artırmaktadır.
- Biçimlendirici değerlendirmenin sınavlarda alınan puanlara olumlu katkı getireceği düşüncesi öğrenciler açısından motive edici olabilmektedir.

Anahtar kelimeler: Biçimlendirici değerlendirme, biçimlendirici dönüt, akran dönütü, öz değerlendirme, fen eğitimi, lisans eğitimi

Submission Checklist Makale Sunumu Kontrol Çizelgesi

Indicate that this submission is ready to be considered by this journal by checking off the following.
Aday makalenin değerlendirilmeye hazır olduğunu aşağıdakilerin her biri ile karşılaştırarak kontrol ediniz.

1	<input type="checkbox"/>	The manuscript is a report of original educational research or a discussion article on research topics in education. Aday makale, eğitim alanıyla doğrudan ilgili bir araştırma ya da tartışma yazısıdır.
2	<input type="checkbox"/>	The submission has not been previously published, nor sent to another journal for consideration. Sunulan çalışma daha önce herhangi bir yerde yayınlanmamıştır, başka bir derginin incelemesinde değildir.
3	<input type="checkbox"/>	In 2010, I have not submitted a manuscript to EJER as an author or co-author for review other than this manuscript. 2010 yılı içerisinde, yazarı olduğum ya da yazarları arasında bulunduğum başka bir çalışma değerlendirilmek üzere EJER'e sunulmamıştır.
4	<input type="checkbox"/>	All authors of this manuscript are subscribers of the EJER and accept to pay 425 TL for the cost of proofreading and 10 loose copies of their articles if the manuscript is accepted. Bu aday makalenin tüm yazarları EJER'e abonedirler ve aday makalenin kabul edilip basılması halinde makalelerinin İngilizce son okuma (proofreading) ve 10 serbest kopya gönderimi nedeniyle oluşan 425 TL maliyeti Anı Yayıncılık'a ödemeyi kabul ederler.
5	<input type="checkbox"/>	The entire manuscript is written in English. Aday makalenin bütünü yazım dili İngilizce'dir.
6	<input type="checkbox"/>	The original manuscript is typed on A4 paper. The margins are 2.5 cm. Aday makale kenar boşlukları 2.5 cm olan A4 kağıda yazılmıştır.
7	<input type="checkbox"/>	Each paragraph is longer than two sentences. Her bir paragraf en az üç cümle içermektedir.
8	<input type="checkbox"/>	The entire manuscript - including quotations, references, author note, content footnotes, figure captions, and all parts of tables - is double-spaced. Aday makalenin tamamı, alıntılar, kaynakça, şekil ve tablo başlıkları da dahil olmak üzere çift aralıklı yazılmıştır.
9	<input type="checkbox"/>	The submission file is in Microsoft Word document file format. 12-point Times New Roman font is used in entire manuscript. Aday makale, tamamında 12 punto Times New Roman yazı tipi kullanılarak hazırlanmış bir Microsoft Word dokümanıdır.
10	<input type="checkbox"/>	The text has had the authors' names removed. If an author is cited, "Author" and year are used in the bibliography and footnotes, instead of author's name, paper title, etc. The author's name has also been removed from the document's Properties, which in MS Word is found under the File menu. Aday makale, yazar adları çıkarılarak sunulmuştur. Eğer yazar kendisine atıfta bulunulursa yazarın adına ve çalışma başlığına yer verilmeyecek, sadece "Author" yazılarak çalışmanın yılı belirtilecektir. Ayrıca, yazarın adı MS Word belge özelliklerinden (Dosya, Özellikler) silinecektir.
11	<input type="checkbox"/>	The title is 10 to 12 words. Aday makalenin başlığı 10-12 sözcük uzunluğundadır.
12	<input type="checkbox"/>	The maximum length of the manuscript-including structured abstract in English, tables, and references is 6000 words. This limitation does not include Turkish extended abstract (750-1000 words) which is placed after the references section. Aday makale, İngilizce abstract, tablolar ve kaynakça vb. tüm öğeler dahil olmak üzere en fazla 6000 sözcüktür. Kaynakça'nın ardından verilen uzun Türkçe özet (750-1000 sözcük) bu sayıya dahil değildir.
13	<input type="checkbox"/>	The article is preceded by English Structured Abstract of not more than 400 words and not less than 300 using five required headings: <i>Problem Statement, Purpose of Study, Methods, Findings and Results, and Conclusions and Recommendations</i> . (These headings may need some adaptation in the case of discussion papers: <i>Background, Purpose of Study, Sources of Evidence, Main Argument, and Conclusions</i>). More information available from (http://www.tandf.co.uk/journals/authors/reabstracts.asp) Yapılandırılmış İngilizce öz 300-400 sözcük uzunluğunda olup, aday makalenin başında yer almakta ve <i>Problem Durumu, Araştırmanın Amacı, Araştırmanın Yöntemi, Araştırmanın Bulguları, Araştırmanın Sonuçları ve Önerileri</i> başlıklarını içermektedir. Bu başlıklar tartışma yazıları için: <i>Çalışmanın Temeli, Çalışmanın Amacı, Kanıt Kaynakları, Ana Tartışma ve Sonuçlar</i> şeklinde olabilir. Daha fazla bilgi için http://www.tandf.co.uk/journals/authors/reabstracts.asp adresine başvurunuz.

14	<input type="checkbox"/>	Following the structured abstract in English four to six keywords are included. Yapılandırılmış İngilizce özden sonra 4-6 anahtar sözcüğe yer verilmiştir.
15	<input type="checkbox"/>	An extended (750-1000 words) Turkish structured abstract is placed following the "References" section using five required headings: <i>Problem Statement, Purpose of Study, Methods, Findings and Results, and Conclusions and Recommendations</i> . (These headings may need some adaptation in the case of discussion papers: <i>Background, Purpose of Study, Sources of Evidence, Main Argument, and Conclusions</i>). More information available from http://www.tandf.co.uk/journals/authors/rereabstracts.asp Kaynakça'dan sonra 750-1000 sözcükten oluşan Türkçe yapılandırılmış öze yer verilmiştir. Türkçe yapılandırılmış öz <i>Problem Durumu, Araştırmanın Amacı, Araştırmanın Yöntemi, Araştırmanın Bulguları, Araştırmanın Sonuçları ve Önerileri</i> başlıklarını içermektedir. Bu başlıklar tartışma yazıları için: <i>Çalışmanın Temeli, Çalışmanın Amacı, Kanıt Kaynakları, Ana Tartışma ve Sonuçlar</i> şeklinde olabilir. Daha fazla bilgi için; http://www.tandf.co.uk/journals/authors/rereabstracts.asp
16	<input type="checkbox"/>	Following the Turkish structured abstract, four to six keywords are included. Uzun Türkçe özetten sonra 4-6 anahtar sözcüğe yer verilmelidir.
17	<input type="checkbox"/>	References are not cited in the structured abstracts in English and in Turkish. İngilizce abstract ve Türkçe öz içerisinde atıfta bulunulmamıştır.
18	<input type="checkbox"/>	The format of headings, tables, figures, citations, references, and other details follow the APA style as described in the Publication Manual of the American Psychological Association, 5th edition, available from http://www.apa.org Aday makalenin başlıkları, tabloları, şekilleri, atıfları, kaynakçası ve diğer özellikleri tamamen APA beşinci baskıda belirtildiği şekildedir.
19	<input type="checkbox"/>	All illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end. Aday makalenin şekilleri ve tabloları metin içerisinde bulunmaları gereken uygun yerlere yerleştirilmiştir. Makale sonunda sunulmamıştır.
20	<input type="checkbox"/>	Citations in the text of the document include the author's surname, the year of publication, and, when there is a specific quote from a source used, a page number where the quote is located in the text. Example: Nothing seemed so certain as the results of the early studies (Tatt, 2001, p. 445). It was precisely this level of apparent certainty, however, which led to a number of subsequent challenges to the techniques used to process the data (Jones & Wayne, 2002, p. 879). There were a number of fairly obvious flaws in the data: consistencies and regularities that seemed most irregular, upon close scrutiny (Aarns, 2003; West, 2003, p. 457). With studies by two authors, always include both author names: (Anderson & Bjorn, 2003) As Anderson and Bjorn (2003) illustrated in their recent study As recently as 2003, a prominent study (Anderson & Bjorn) illustrated When a study has 3, 4, or 5 authors, include the names of all the authors the first time the work is cited: (Anderson, Myers, Wilkes, & Matthews, 2003) For all subsequent citations of this work, use "et al.": (Anderson et al., 2003) When a work has 6 or more authors, use et al.: (Bell et al., 2003) For unsigned works, include the title, enclosed in parentheses. Put quotation marks for short work titles, and italicize the titles of reports, books, and other significant works: ("Recent Developments," 2004) (Dictionary of Tetrathalocigistic Diseases, 2004) Metin içindeki atıfları üstte verilen örneklere uygundur.
21	<input type="checkbox"/>	Three levels of headings are used: Level 1, Level 3 and Level 4. The headings are formatted as follows: Centered Uppercase and Lowercase Heading (Level 1)

		<p><i>Flush Left, Italicized, Uppercase and Lowercase Side Heading (Level 3)</i> <i>Indented, italicized, lowercase paragraph heading ending with a period. Start writing after the period (Level 4).</i></p> <p>Aday makale içerisinde üç farklı düzey başlık kullanılmıştır. Düzey 1, Düzey 2, Düzey 3. Başlıklar bu düzeylere uygun olarak aşağıdaki şekilde biçimlendirilmiştir: Ortalı ve Her Sözcüğün İlk Harfi Büyük Yazılmış Başlık (Düzey 1) <i>Tam Sola Dayalı, İtali ve Her Sözcüğün İlk Harfi Büyük Yazılmış Başlık (Düzey 3)</i> İçeriden, itatik, tamamı küçük harflerle yazılmış ve nokta ile bitten başlık. Noktadan sonra normal metin yazımına devam edilmeli (Düzey 4).</p>
22	<input type="checkbox"/>	<p>References are listed in alphabetical order. Each listed reference is cited in text, and each text citation is listed in the References. Basic formats are as follows: Haag, L., & Stern, E. (2003). In search of the benefits of learning Latin. <i>Journal of Educational Psychology, 95</i>, 174-178. Bollen, K. A. (1989). <i>Structural equations with latent variables</i>. New York: Wiley. Johnson, D. W., & Johnson, R. T. (1990). Cooperative learning and achievement. In S. Sharan (Ed.), <i>Cooperative learning: Theory and research</i> (pp. 173-202). New York: Praeger.</p> <p>More information is available from: http://citationonline.net/CitationHelp/csg04-manuscripts-apa.htm#references</p> <p>Kaynakçanın yazımı üstte verilen örneklere uygundur.</p>
23	<input type="checkbox"/>	<p>Order of the main parts in the manuscript is as follows: Main title in English (max. 12 words) Structured abstract (min. 300- max.400 words length) Keywords (in English, min. four-max. six) Main text References Main Title in Turkish (max. 12 words) Extended structured abstract (min.750-max.1000 words length in Turkish) Keywords (in Turkish, min. four-max. six)</p> <p>Aday makaleyi oluşturan ana öğeler aşağıdaki düzendedir: İngilizce Ana Başlık (En fazla 12 sözcük) Yapılandırılmış İngilizce Abstract (En az 300, en fazla 400 sözcük) Anahtar Sözcükler (İngilizce, en az dört, en fazla altı) Ana Metin Kaynakça Türkçe Ana Başlık (En fazla 12 sözcük) Yapılandırılmış Türkçe Öz (En az 750, en fazla 1000 sözcük) Anahtar Sözcükler (Türkçe, en az dört, en fazla altı)</p>
24	<input type="checkbox"/>	<p>The manuscript and this checklist form are submitted electronically to ejer.editor@gmail.com. Aday makale ve bu kontrol çizelgesi elektronik olarak ejer.editor@gmail.com adresine gönderilmiştir.</p>

ABONELİK

Yıllık 4 Sayı İçin:

Kişi Aboneliği.....: 70

TL

Kurum Aboneliği.....: 200 TL

Yurtdışı Aboneliği.....: 250 \$

Öğrenci Aboneliği.....: 60 TL

Tek Sayılar İçin:

Kişiler: 25 TL

Kurumlar: 75 TL

ÖDEME

Anı Yayıncılık Posta Çeki Hesap No: 1911545 veya

İş Bankası Kızılay Şubesi, 4214-0962561 numaralı hesaplar.