

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: 9 Issue: 36 / 2009
Üç Ayda Bir Yayınlanan Hakemli Dergi, Yıl: 9 Sayı: 36 / 2009

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

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

INTERNATIONAL EDITORIAL BOARD / Uluslararası Editörler Kurulu

Adnan Boyacı, Anadolu University, Eskişehir, TURKEY
Adnan Erkuş, Mersin University, Ankara, TURKEY
Anita Pipere, Daugavpils University, LATVIA
Aslı Özgün Koca, Wayne State University, USA
Beth Ashforth, Northumbria University, Newcastle, UK
Danny Wyffels, KATHO University, Kortrijk, BELGIUM
Jordanescu Eugen, Lucian Blaga University of Sibiu, ROMANIA
İbrahim Yıldırım, Hacettepe University, Ankara, TURKEY
İsmail Ö. Zembat, United Arab Emirates University, Abu Dhabi, UAE
Mehmet Buldu, United Arab Emirates University, Abu Dhabi, UAE
Murat Özbay, Gazi University, Ankara, TURKEY
Selahattin Gelbal, Hacettepe University, Ankara, TURKEY
Sven Persson, Malmö University, Malmö, SWEDEN
Tuncay Ergene, Hacettepe University, Ankara, TURKEY
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, Kızılırmak 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: 15.07.2009
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: ANI Publishing / Anı Yayıncılık
Typography / Dizgi: Ali Ertuğrul
The ideas published in the journal belong to the authors.
Dergide yayınlanan 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.
© 2008 ANI Publishing. All rights reserved. © 2008 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

Editorial: Bridging the Research Divide between Counseling and Related Fields
Senel Poyrazli, Michael J. Stevens ----- 1-6

The Psychometric Qualities of the Turkish Version of Emotional Abuse Questionnaire
Günnur Karakurt, Bengü Ergüner-Tekinalp, Şerife Terzi -----7-22

Counselors and Administrators: The Collaborative Alliance in Three Countries
Fidan Korkut-Owen, Dean W. Owen, Victor Ballesteros -----23-38

A Study of Graffiti in Teacher Education
S.Nihat Şad, Mustafa Kutlu -----39-56

An Investigation of University Students' Critical Thinking Disposition and Perceived Problem Solving Skills
Songül Tümkaya, Birsal Aybek, Habibe Aldağ -----57-74

Investigation of University Students' Depression
A.Aykut Ceyhan, Esra Ceyhan, Yıldız Kurtyılmaz -----75-90

An Application of Digital Portfolio with the Peer, Self and Instructor Assessments in Art Education
Ayhan Dikici -----91-108

The Effects of Computer Animations and Cooperative Learning Methods in Micro, Macro and Symbolic Level Learning of States of Matter
Kemal Doymuş, Ümit Şimşek, Ataman Karaçöp -----109-128

Examining the Relationship between Epistemic Curiosity and Achievement Goals
Altay Eren -----129-144

Does Being Leisure-Oriented Preclude a Person's Commitment to Work?
Şöheyda Göktürk -----145-160

The Barriers Hindering Female Teachers From Becoming School Principals (Sample from Mersin Province-Turkey)
Yusuf İnandı -----161-178

Reliability and Validity Study of Turkish Form of the Psychological Distress Scale
Hülya M. Ünal Karagüven -----179-192

The Effects of an Emotional Intelligence Skills Training Program on the Emotional Intelligence Levels of Turkish University Students
T.Fikret Karahan, B. Murat Yalçın -----193-208

Bullying and School Climate from the Aspects of the Students and Teachers
Hülya Kartal, Asude Bilgin -----209-226

Development of a Two-Tier Diagnostic Test to Determine Students' Understanding of Concepts in Genetics
Didem Kılıç -----227-244

Views of Primary School Administrators, Teachers and Parents on Parent Involvement in Turkey
Ali Sabancı -----245-262

A Study on Developing a Self Efficacy Scale towards Science and Technology
Nilgün Tatar, Eylem Yıldız, Ercan Akpınar, Ömer Ergin -----263-280

REVIEWERS of the 36th ISSUE
36. Sayı Hakemleri

Reviewers of the Special Issue

Amy Lyubchik
Ayşe Ciftci
Carol Hoare
Chalmer Thompson
Deniz Canel Cinarbas
Gina Brelford
Guler Boyraz
Jeeseon Park
Johanna Nilsson
Katherine Ziff
Mark Leach
Mathilde Salmberg
Metin Ozdemir
Oksana Yakusko
Victoria A. Shivy
Wade Pickren
William Chris Briddick

Reviewers of the 36th Issue

Ahmet Ok
Ali Eryılmaz
Ata Tezbaşaran
Ayhan Demir
Ayşe Balcı
Bayram Çetin
Behzat Bektaşlı
Berrin Burgaz
Burcu Akbayrak
Elif Üstün
Ercan Kiraz
Erdinç Çakıroğlu
Erdinç Duru
Fidan Korkut
Hasan Hüseyin Aksoy
Hülya Kelecioğlu
İbrahim Yıldırım
İnayet Pehlivan
Kazım Artut
Leman Tarhan
M. Hakan Gündoğdu
Mehmet Buldu
Mehtap Çakan
Meral Atıcı
Meral Hakverdi
Muammer Çalık
Naciye Aksoy Kuş
Nalan Yılmaz
Nuri Doğan
Ömer Kutlu
Özgül Yılmaz Tüzün
Robert Sabol
Selahattin Gelbal
Serkan Yılmaz
Şakir Çınkır
Şenel Poyrazlı
Tolga Arıca
Yusuf Badavan
Yüksel Kavak

FROM THE GUEST EDITORS

Bridging the Research Divide between Counseling and Related Fields

Senel Poyrazli*
Michael J. Stevens**

In this special section of *Eurasian Journal of Educational Research* we introduce four articles based on the collaborative work of scholars whose training and expertise reflect diverse disciplinary traditions. Over one year ago, the editorial board of the Journal decided to publish special issues focusing on topics of interest and relevance to the readership. One of these special issues was to center on counseling and closely aligned fields. Contemporary topics within counseling were identified and reviewed, after which the associate editors responsible for counseling-related journal contributions determined that this special issue would feature interdisciplinary research in the areas of guidance and counseling, counselor education, and counseling psychology. The co-editors of this special issue sought interdisciplinary research that involved quantitative, qualitative, and/or mixed methodologies, research with an applied focus, and research that examined and/or proposed social policies or interventions. The co-editors stipulated that submissions needed to be co-authored, with at least one author trained in guidance and counseling, counselor education, or counseling psychology and one or more authors from a closely aligned field (e.g., clinical psychology, educational psychology, educational leadership, elementary education).

Twenty-two manuscripts were submitted for publication in this special issue. The co-editors recruited internationally recognized authorities to serve as reviewers. The editorial decision to accept or reject a manuscript rested on the evaluations and recommendations provided by reviewers. After several rounds of reviews, four manuscripts were accepted for publication. These articles appear in this special section of the Journal.

As co-editors, we would like to use this special section as an opportunity to inform the Journal's readership about interdisciplinary research, particularly its contemporary relevance and potential value. Interdisciplinary research is customarily defined as research that is conducted by experts who represent different

* PhD, The Pennsylvania State University - Harrisburg, USA

** PhD, DHC, Illinois State University, USA The Lucian Blaga University, Romania

scientific fields and as research that incorporates perspectives from more than one discipline (Aboelela et al., 2006). Although the importance of interdisciplinary research has been discussed and promoted for decades (see Sherif & Sherif, 1969, as an example), it is only recently that interdisciplinary research has gained credibility in certain fields as an approach to studying complex, novel, and/or pressing phenomena. Many academic departments have begun to promote interdisciplinary scholarship, even restructuring organizationally in order to facilitate such scholarship. Grant-funding entities (e.g., the European Commission) increasingly require that investigators infuse an interdisciplinary approach throughout their proposals. One explanation for the growing emphasis on interdisciplinary research is that, as scholars have advanced the scientific knowledge in their own disciplines, they "have become increasingly aware of the need to link disciplinary fields to more fully answer critical questions, or to facilitate application of knowledge in a specific area" (Aboelela et al., p. 330). This awareness stems from concerns about the capacity of any one disciplinary specialty to meet the challenges of a global era. These concerns have prompted a critical examination of the assumptions underlying delimited fields in the social sciences, including those on which research in counseling are grounded. More generally, critics have judged the social sciences to be fragmented because of their reductionistic orientation, in which complex and contextualized phenomena are isolated in laboratories and dismantled in experimental manipulations that rob them of meaning. As a result, social scientists and their students, social science departments, and professional organizations in the social sciences are becoming more receptive and willing to adopt innovative approaches in the study of vexing phenomena that cross artificially segmented disciplinary boundaries and that offer the promise of yielding more externally valid and useful findings (Stevens & Wedding, 2004).

Unlike disciplines and specialties such as social work and community psychology, guidance and counseling, counselor education, and counseling psychology have only recently begun to encourage interdisciplinary scholarship. Although the benefits of collaboration among colleagues within the same field have long been promoted, the potential value of joining forces with closely aligned colleagues has until recently remained unrecognized and, at times, even been diminished (Palmer, 2004). Counselors and counseling psychologists have gradually come to realize that in order to serve their clients well, to design and implement the most effective remedial and preventative methods, to win competitive grants, and to conduct meaningful research and publish in high-impact outlets, they need to "collaborate with other professionals and pool ... [their] talents and expertise" (Borges, McNally, Maguire, Werth, & Britton, 2008, p. 128). As a result, many counselors and counseling psychologists have sought to establish collaborative relationships with scientists and practitioners in closely aligned disciplines in order to advance their research and applied agendas. A compelling illustration can be found in interdisciplinary efforts to prevent child abuse. Counselors, counseling psychologists, educators, physicians, social workers, and other allied professionals have worked in tandem to develop new and more effective approaches to abuse prevention (Palmer, 2004). Another example can be found in the treatment of eating

disorders. Harris (1991) describes a research project in which faculty and students first identified bulimia nervosa as a major health issue in their university community. Next, they formed an interdisciplinary team of counseling psychologists, dietitians, marketing professionals, and physicians. Together, they designed and promoted an innovative intervention program to combat bulimia nervosa. Dietitians devised a means whereby food intake could be monitored via computers, marketing specialists worked to “sell” the intervention, and physicians served as medical consultants and also referred individuals to the program.

Although we provide a list of references that offer useful information for those who seek to establish or increase interdisciplinary collaboration in their research programs, we are obliged to note that interdisciplinary research may not always be a necessary, let alone optimal, approach to conducting research. In fact, based on their experience, Crow, Levine, and Nager (1992) identify several challenges that researchers might face when pursuing interdisciplinary projects. They explain that researchers’ background and training may lead them to favor competing methodologies for studying the same phenomenon. They indicate that this is especially likely to occur when contemplating the merits of applying qualitative versus quantitative methods to the investigation of a particular phenomenon in a particular context. They propose that ongoing dialogue about different research paradigms and their underlying assumptions, coupled with deliberate efforts to nurture great acceptance of diverse methods of inquiry, can establish norms for productive communication within an interdisciplinary research team and prevent the emergence of conflict that could stymie the research enterprise. Crow et al. also warn that a dominant, monopolizing voice from one discipline can inhibit other disciplinary perspectives from being heard, thereby weakening the research design and the cohesiveness of the research team. They propose that maintaining awareness of patterns of communication as well as respect for diverse input on how to approach the research project can yield a stronger, integrated study. Notwithstanding the challenges they encountered, Crow et al. believe that the inclusion of researchers from other disciplines enhanced their own research.

Takooshian and Stevens (2001) remark on the dearth of publications on interdisciplinary collaboration. Although interdisciplinary collaboration can be accomplished in various ways, the initiator must first evaluate his or her expertise and resources as well as the complementary strengths and weaknesses of potential collaborators. Once collaborators join an interdisciplinary research team, it is vital to delineate the objectives of and process required by the research project. With respect to the personal dimension of interdisciplinary collaboration, one must understand the nuances of initiating and nurturing interdisciplinary relationships. For example, in order to avoid potentially disruptive power differentials, each team member should have a significant role in the research project, ideally one that is established by consensus. At times, it may be necessary to assume a leadership role when focusing the team on important tasks, motivating a high level of engagement, and resolving misunderstandings.

As co-editors of this special section, we both agree with the view that interdisciplinary collaboration can enrich a research program and advance one's field. Although we are counseling psychologists by training and have collaborated with other counseling psychologists, we also have joined with counselor educators, clinical psychologists, developmental psychologists, social psychologists, anthropologists, sociologists, social workers, nurses, educators, linguists, and police to name a few. These collaborations have enabled us to enhance the sensitivity and depth with which we studied various phenomena of interest. As a result, we believe that our research generated knowledge that otherwise might not have emerged from a more traditional, single disciplinary approach.

Establishing Fruitful Interdisciplinary Collaboration

As co-editors of this special section, we would like to offer suggestions for those who may wish to initiate or strengthen their interdisciplinary research collaboration.

1. Read books and journal articles written by scholars in fields other than your own, including publications on diverse methodologies appropriate to interdisciplinary research. Among many possibilities are the books, *Critical Psychology: Voices for Change* (Sloan, 2000), *Crossing Boundaries: Knowledge, Disciplinarity and Interdisciplinarity* (Klein, 1996), *Handbook of Culture, Therapy, and Healing* (Gielen, Fish, & Draguns, 2004), and *Qualitative Research in Psychology: Expanding Perspectives in Methodology and Design* (Camic, Rhodes, & Yardley, 2003). Relevant journal articles include, "Are Three Heads Better than One? Reflections on Doing Collaborative Interdisciplinary Research" (Crow et al., 1992), the special issue on "International Perspectives in Counseling Psychology" in *Applied Psychology: An International Review* (Leong & Savickas, 2007), "Anglo-North American Qualitative Counseling and Psychotherapy Research" (Rennie, 2004), and "Quality Control in Qualitative Research" (Stiles, 1993). The e-mail addresses of potential collaborators often appear at the end of their publications or can be found on the Internet.

2. If you teach, incorporate research from closely aligned disciplines in your classes (Woods, 2007). Not only will you become familiar with research in related disciplines, but you also will be able to map out potential future collaboration and demonstrate to your students that the same phenomenon can be approached meaningfully from a variety of diverse theoretical and methodological perspectives. For example, exposure to research on student-teachers' perceptions of barriers to student learning could enable you and your students to recognize additional factors that contribute to academic achievement or failure.

3. If you have yet to complete an advanced degree or are required to enroll in continuing education, consider registering for a course in a closely aligned field. For instance, a sociology course on social structures may heighten your understanding and appreciation of how different social contexts can affect a person's psychological experiences.

4. If you are an academic, ask your department chair or college dean to introduce you to scholars in other disciplines who share your research interests. Consider inviting a university administrator to join you in developing an on-campus interdisciplinary directory based on the research interests and programs of colleagues from various departments and colleges.

5. Expand your professional network by adding colleagues from other fields. Inaugurate an online interdisciplinary team to conduct a research project of mutual interest that is regional, national, or international in scope. Consider posting a notice on the web site, listserv, or newsletter of a professional organization in your discipline and/or closely aligned discipline. Or, join an allied professional organization or division within your professional organization that has a special interest group on interdisciplinary collaboration.

6. While at conferences, attend research presentations in areas similar to your own, but within another division's program. This will require that you review the conference program in advance and identify relevant presentations in closely aligned divisions (e.g., programming in clinical psychology, school psychology, psychotherapy, educational psychology, international psychology).

7. After attending the presentations just described, introduce yourself to the presenter(s), briefly state your areas of research interest and expertise, and highlight the reasons you attended the presentation. Exchange business cards and follow-up with an e-mail to inquire about the possibilities of interdisciplinary collaboration and to establish ongoing discourse. Key questions concerning interdisciplinary research await answers (Stevens & Gielen, 2007). These questions that can be meaningfully explored through an ongoing global, electronic conversation. For example, "Under what conditions are reductionistic, normative, and mixed conceptual models most likely to enrich understanding and provide guidance for research and practice?" "To what extent can interdisciplinary collaboration serve to incorporate such dimensions as individualism-collectivism, science-folklore, and secularism-religion into theory, research, and practice." "When are quantitative, qualitative, and mixed-method investigatory approaches most appropriately applied?"

8. New avenues for interdisciplinary research have been paved by advancements in telecommunications. By utilizing various Internet-based technologies, scholars in guidance and counseling, counselor education, counseling psychology, and closely aligned disciplines are now able to archive and share their data and research findings with the aim of enhancing the coherence and utility of knowledge. Furthermore, there are many pressing social issues, some of which are addressed in this special section, whose resolution would likely accelerate if examined in an interdisciplinary fashion.

References

- Aboelela, S. W., Larson, E., Bakken, S., Carrasquillo, O., Formicola, A., Glied, S. A., Haas, J., & Gebbie, K. M. (2007). Defining interdisciplinary research: Conclusions from a critical review of the literature. *Health Services Research, 42*, 329-346.
- Borges, N. J., McNally, C. J., Maguire, C. P., Werth, J. L., & Britton, P. J. (2008). Work, health, diversity and social justice: Expanding and extending the discussion. *The Counseling Psychologist, 36*, 127-131.

- Camic, P. M., Rhodes, J. E., & Yardley, L. (Eds.). (2003). *Qualitative research in psychology: Expanding perspective in methodology and design*. Washington, DC: American Psychological Association.
- Crow, G., Levine, L., & Nager, N. (1992). Are three heads better than one? Reflections on doing collaborative interdisciplinary research. *American Educational Research Journal*, 29, 737-753.
- Gielen, U. P., Fish, J., & Draguns, J. G. (Eds.). (2004). *Handbook of culture, therapy, and healing*. Mahwah, NJ: Erlbaum.
- Harris, J. K. (1991). Evolution of a health focus: Old values – new practices in counseling psychology education and training. *The Counseling Psychologist*, 3, 365-375.
- Klein, J. T. (1996). *Crossing boundaries: Knowledge, disciplinarity and interdisciplinarity*. Charlottesville, VA: University Press of Virginia.
- Leong, F. T. L., & Savicaks, M. L. (Eds.). (2007). International perspectives on counseling psychology [Special issue]. *Applied Psychology: An International Review*, 56(1).
- Palmer, L. K. (2004). The call to social justice: A multidiscipline agenda. *The Counseling Psychologist*, 32, 879-885.
- Rennie, D. L. (2004). Anglo-North American qualitative counseling and psychotherapy research. *Psychotherapy Research*, 14, 37-55.
- Sherif, M., & Sherif, C. W. (Eds.). (1969). *Interdisciplinary relationships in the social sciences*. Chicago: Aldine Publishing.
- Sloan, T. S. (Ed.). (2000). *Critical psychology: Voices for change*. New York: St. Martin's Press.
- Stevens, M. J., & Gielen, U. P. (2007). Epilogue. In M. J. Stevens & U. P. Gielen (Eds.), *Toward a global psychology: Theory, research, intervention, and pedagogy* (pp. 391-398). Mahwah, NJ: Erlbaum.
- Stevens, M. J., & Wedding, D. (2004). International psychology: A synthesis. In M. J. Stevens & D. Wedding (Eds.), *Handbook of international psychology* (pp. 481-500). New York: Brunner-Routledge.
- Stiles, W. B. (1993). Quality control in qualitative research. *Clinical Psychology Review*, 13, 593-615.
- Takooshian, H., & Stevens, M. J. (2001). Collaborating on cross-national survey research: Why and how. *International Psychology Reporter*, 5(3-4), 13-15.
- Woods, C. (2007). Researching and developing interdisciplinary teaching: Towards a conceptual framework for classroom communication. *Higher Education*, 54, 853-866.

The Psychometric Qualities of the Turkish Version of Emotional Abuse Questionnaire

Günnur Karakurt*
Bengü Ergüner-Tekinalp**
Şerife Terzi***

Suggested Citation:

Karakurt, G., Ergüner-Tekinalp, B., & Terzi, Ş. (2009). The psychometric qualities of the turkish version of emotional abuse questionnaire, *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 7-22.

Abstract

Problem Statement: Emotional abuse in intimate relationships is a serious problem that might have devastating consequences for individuals who are subjected to it. Therefore, assessment of emotional abuse is critical for the identification, recognition, and study of this important problem in romantic relationships.

Purpose of Study: The purpose of this study is to translate and validate the Emotional Abuse Questionnaire (EAQ) in Turkish culture.

Methods: Participants of this study included 95 college students. The EAQ was translated from English to Turkish with forward and reverse translation methods. In order to test the factor structure of each EAQ subscale (isolation, degradation, sexual abuse and property damage), principle component analysis was performed. The Dyadic Adjustment Scale was used to establish the validity of the criteria used. In order to control for social desirability, the short-version of the Marlowe-Crowne Social Desirability Scale was used.

Results: Examination of scree plot and initial analyses and communality scores for the EAG isolation and degradation subscales indicated that one-factor solutions were appropriate. The results yielded that a seven-item, one-factor solution best represents the sexual abuse subscale. Finally, a seven-item, one-factor solution emerged for the property damage subscale. Internal consistency and test-retest reliabilities suggest that the Turkish version of the Emotional Abuse Questionnaire has high internal coherence and high temporal stability. These reliability scores are comparable to those of the original version. The validity of the criteria for the Turkish EAQ was evaluated against the Dyadic Adjustment Scale. As hypothesized, Turkish

* PhD, Texas Tech University, Department of Applied & Professional Studies, gunnur.karakurt@ttu.edu.

** PhD, Drake University, Leadership Counseling and Adult Development Department, bengü.tekinalp@drake.edu.

*** PhD, Gazi University Faculty of Education, serife@gazi.edu.tr.

EAQ scores were negatively correlated with scores on the Turkish Dyadic Adjustment Scale.

Conclusion and Recommendations: The results of this study indicated that the EAQ's isolation, degradation, and property damage subscales are applicable to the Turkish culture. The sexual abuse subscale, on the other hand, did not demonstrate sufficient validity to justify its use in Turkish culture among university students. Consequently, the validity of this subscale should be further studied in diverse samples, including married couples.

Keywords: emotional abuse, Emotional Abuse Questionnaire, Dyadic Adjustment Scale, romantic relationships.

Intimate partner abuse is extensive throughout the world and is often severe (Garcia-Moreno, Jansen, Ellsberg, Heise, & Watts, 2006). Abuse in intimate relationships can occur in physical or nonphysical forms. Acts such as hitting, kicking, and choking are among the physical manifestations of abuse, while emotional abuse includes acts such as verbal assault, dominance, control, isolation, ridicule, use of intimate knowledge for degradation, or threats of violence directed toward an individual (Follingstad, Rutledge, Berg, Hause, & Polek, 1990; Goodman, Koss, & Felipe-Russo, 1993; Hamby, 1996; Katz & Arias, 1999; Marshall, 1999; Stets, 1990; Straus, 1979; Tolman, 1999). Many researchers use the term emotional abuse interchangeably with psychological abuse, verbal abuse, psychological maltreatment, and psychological aggression. Studies on intimate partner violence often focus on physical abuse, overlooking the importance of emotional abuse in intimate relationships. However, recent research demonstrated that emotional abuse is a unique form of partner abuse with its own etiology and consequences. All couples in romantic relationships get angry with each other. Even contented couples may confront one another or act in a way that gives the impression of emotional abuse in the heat of an argument. However, these interactions do not intend to intimidate, subjugate, or control another human being (Jacobson & Gottman, 1998). The key elements are coercive and aversive behaviors that aim to inflict emotional harm. Abusers often target the recipient's emotional well-being or sense of self. Furthermore, emotional abuse consolidates power and preserves fear by exploiting past and present violent episodes and constant threat of future violence (Jacobson & Gottman, 1998). This wide definition of emotional abuse can manifest itself in an array of behaviors spanning the continuum from occasional mildly coercive behaviors to the most extreme forms of violence (Romero, 1985).

Previous studies documented that recipients of emotional abuse suffer negative health consequences (O'Leary, 1999). Emotional abuse is frequently associated with the feeling of fear, increased dependency, and damage to self-concept (Murphy & Hoover, 1999). Researchers also demonstrated a high correlation between physical abuse and emotional abuse (Gondolf, Heckert, & Kimmel, 2002; Wilson, Johnson, & Daly, 1995). Furthermore, longitudinal studies reveal that psychological aggression predicts the beginning and occurrence of later physical abuse (Leonard & Senchak, 1996). Such studies illuminate developmental paths associated with intimate partner

violence. Specifically, the early emergence of verbal abuse in a relationship was found to be a risk factor for physical abuse later in a relationship (Leonard & Senchak, 1996; Schumacher & Leonard, 2005). However, other studies showed that emotional abuse can occur without physical abuse (Joiner, 2008). These previous research findings draw attention to emotional abuse and motivate the study of emotional abuse as an independent construct.

In the past 15 years, emotional abuse has been investigated in various cultures, including Denmark, Finland, Sweden, Iceland, Norway (Wijma et al., 2003), Canada (Pejic, 2005), and Turkey (Çelik & Bayraktar, 2004). Some of these studies focused on emotional abuse in the context of health systems in a number of different cultures. However, more research is needed to understand emotional abuse and its consequences in the context of intimate partner relationships in different cultures.

Turkey is one of the countries that presents complex cultural and historical influences from Europe, the Mediterranean, the Middle East, and Asia. It is a secular country with a Muslim majority population. Turkey's population is approximately 70.5 million with an annual growth rate of 1.04 percent, with approximately half of the population younger than 25 years old (State Institute of Statistics (SIS), 2008). In Turkey, both "modern" and "traditional" lifestyles co-exist simultaneously (Atalay, Konaş, Beyazıt, & Mağdenoğlu, 1992; Kağıtçıbaşı, 2000). Since the 1950s, Turkey has undergone an extensive process of urbanization which has led to a higher rate of the population (59 percent) living in urban areas, while 41 percent of the population lives in rural areas (SIS, 1990). However, strong family relationships and their influences in the formation of values, attitudes, aspirations, and goals are still observed in current Turkish culture (Kağıtçıbaşı, 2000; Kağıtçıbaşı, 2003). Many Turkish women are economically dependent and, compared to men, are less likely to be involved in the work force. According to United Nations Economic Commission for Europe (UNECE), less than 30 percent of the total labor force in Turkey consists of women (UNECE, 2001).

In Turkish culture, patriarchal values and codes of honor regulate the social life, particularly in rural areas; special attention is given to the chastity of women (Sunar & Fişek, 2005). Akın and Özeydın (2005) investigated men's attitudes toward women's status in the family and society as well as violence in the home. Half of the men stated that "the important decisions in the family should be made by the male" (p. 201) and "a woman should not argue with her husband even if she does not share the same view with him" (p. 201). Nearly one in four married men agreed with the statement "it is better for a male child to be educated than a female child" (p. 201). One third of men agreed to the statement that "men are usually wiser than women" (p. 201) and nearly half of men had an unfavorable opinion about women's status at home and in the family. Most of the men agreed that if their wives talk with other men and if they argue with their husbands and neglect the care of their children, husbands have the right to beat their wives. One third of the men had a supportive attitude toward physical abuse. Men also placed women in an inferior power position compared to the adult male members of the household (Akın & Özeydın, 2005).

Öngen (2006) explored attitudes toward women in a university sample. Overall, the findings indicated that female university students in Turkey were more liberal than males. Women articulated more liberal views on the vocational, educational, and intellectual issues in marital relations. Female university students do not believe they are subordinate to men, especially in terms of career issues. Moreover, women do not perceive housework and child rearing as solely female responsibilities or obligations. Women also reported more liberal attitudes than men for drinking, swearing and dirty jokes. However, they reported more traditional views on these issues in comparison to others. On the other hand, men held more traditional attitudes; they expressed that women should be responsible for child care and household management. Men agreed that social norms biased in their favor should be maintained and women's role in society should not change. Men reported more liberal views on dating, etiquette, and courtship. Both men and women reported conservative attitudes toward premarital sex (Öngen, 2006). As men and women tend to differ in their attitudes toward traditional gender roles, it is important to explore the interplay between these attitude differences and relationship dynamics, including those that relate to emotional abuse.

In order to study emotional abuse in a particular culture, researchers need well-established measures that are applicable in their cultural context. Various measures have been developed to understand emotional abuse. One of the most commonly used measures is the Emotional Abuse Questionnaire developed by Jacobson and Gottman (1998). The aim of the present study is the adaptation of this questionnaire into the Turkish language and its validation in the context of Turkish culture. In particular, this study aims to assess the reliability and factorial structures of the Turkish translation of the Emotional Abuse Questionnaire when used with Turkish samples.

Method

Participants

The participants of this study included 95 college students in Turkey who were currently in a romantic relationship. Of the participants, 82 were female (86 percent) and 12 (14 percent) were male. The average age of the participants was 20.58 ($SD=1.94$) and ranged from 18 to 28 years old. The mean duration of being involved in their relationship was 11 months ($SD = 17.86$), ranging from one month to five years.

Instruments

Emotional Abuse Questionnaire (EAQ). The EAQ is a self-administered test that assesses different aspects of emotional abuse in intimate relationships (Jacobson & Gottman, 1998). The EAQ was developed as a part of Jacobson and Gottman's (1998) larger study on violent marriages. Data for this study was collected from 201 couples in the Seattle, Washington, area. Couples were divided into different groups (battering couples, couples that exhibited some violence, dissatisfied non-violent

couples, and happily married couples), based on multiple measures such as the Conflict Tactic Scale and the Specific Affect Coding System (SPAFF). This grouping allowed researchers to compare violent couples to non-violent couples, and to unhappily married couples, to account for marital distress. Husbands and wives were interviewed individually about their violent and non-violent arguments in the past. They were also followed up for two weeks following initial data collection and were queried about the worst argument they had during the period between the two waves. In order to assess emotional arousal physiologically, couples were videotaped while they were discussing an issue of disagreement in their marriage. Participants' bodily reactions such as heart rate, blood flow, and sweating were monitored by researchers with electronic sensors. These measures were compared with baseline levels. Later, researchers transcribed the videotapes and analyzed the data using a detailed system that classified visual gaze, vocalization, attempts to influence one another, and emotional expressions. Discussions, behaviors, and physical responses were synchronized to understand the violent interactions. In order to understand the changes over time, couples were studied again two years later, investigating the differences in their arguments and level of emotional abuse. Results of this eight-year extensive study indicated four distinct categories for emotional abuse—destruction of pets and property, sexual coercion, attempts to isolate, and degradation. These different categories were also supported with their extensive interviews (Jacobson & Gottman, 1998). The subscales of the instrument were constructed based on these categories and the items were selected based on the themes associated with these categories.

The EAQ has 66 items ranked by participants on a Likert-type scale from 1 (*Never*) to 4 (*Always*). It has four separate subscales: isolation, degradation, sexual abuse and property damage. Internal consistency for the subscales are .92, .94, .72, and .82, respectively. The isolation subscale is composed of 24 items, including "My partner tries to control whom I spend time with" and "My partner has disabled the car." The degradation subscale has 28 items, including "My partner humiliates me in front of others" and "My partner ridicules me." The sexual abuse subscale has seven items including "My partner makes me engage in sexual practices I consider perverse." The property damage subscale has seven items including "My partner intentionally damages things that I care about." Responses to individual items are added to form the total score; higher scores indicate higher emotional abuse. Negatively worded items were reverse coded.

Dyadic Adjustment Scale (DAS; Spanier, 1976). The DAS was used to establish the criterion validity of the Emotional Abuse Questionnaire. The DAS assesses the quality of the relationship in intimate romantic relationships; the existence of emotional abuse would lower the quality of an intimate relationship. The Dyadic Adjustment Scale was chosen to test criteria-related validity. We hypothesized that the Dyadic Adjustment Scale would negatively correlate with the Emotional Abuse Questionnaire and the presence of emotional abuse in intimate relationships would decrease the perception of the quality of the relationship. The Dyadic Adjustment Scale was translated and adapted into Turkish by Fıfılođlu and Demir (2000) by

using one directional translation method (Hambleton & Bollwark, 1991). The 32 self-report items of the DAS were designed to assess the perceived quality of the relationship. The instrument has four subscales, measuring four distinct aspects of a relationship. The subscales are dyadic satisfaction, dyadic cohesion, dyadic consensus, and affectional expression. Participants use a five- to seven-point Likert scale ranging from *Always agree* to *Always disagree* or *All the time* to *Never*. Scale scores are obtained by adding the items specific to each subscale. Responses to individual items are added to form the total score. Higher scores indicate a more positive perception of the quality of the relationship. Internal consistency reliabilities (Cronbach's alphas) were reported as .96 for the entire scale and the subscale reliabilities ranged from .73 to .94, suggesting high reliability. The adapted scale showed high reliability and good criteria validity in the Turkish cultural context. The scale also was reported to have good content validity (Fişiloğlu & Demir, 2000). For the Turkish version of the scale, the Cronbach alpha was reported as .92 and split-half reliability as .86 (Fişiloğlu & Demir, 2000). In order to use the questionnaire with dating couples, minor changes in the items were made. The word "spouse" was replaced with "the person I am dating" and the item "Have you ever regretted that you are married to your spouse?" was replaced with the item "Have you ever regretted that you are dating this person?"

Marlowe-Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1960). In order to test for social desirability bias, the widely used MCSDS short version was used. The short version of the scale was reported to be as internally consistent as the long version (Fraboni & Cooper, 1989; Zook & Sipps, 1985). The short version of MCSDS was translated and adapted into Turkish by Ural and Özbirecikli (2006). Forward and back translation method was used and the Turkish version of the scale yielded high reliability and acceptable validity. In the final short version of the Turkish MCSDS, participants indicate agreement or disagreement using a six-point Likert scale (1=*Strongly disagree* to 6=*Strongly agree*) to rate socially desirable behaviors. A total score was obtained by adding the item scores. Higher scores indicate a greater tendency toward social desirability. The internal consistency reliability of the Turkish short version of MCSDS was reported as .78 (Ural & Özbirecikli, 2006).

Procedure

The participants were college students who were currently in a romantic relationship. The participants were recruited by purposive sampling from different major universities in Ankara, Turkey. During recruitment the third author gave verbal instructions on the purpose of the study, how to fill out the instruments, and participant rights that the participation is voluntary and they can withdraw any time without negative consequences. Those participants who chose to participate were asked not to write their names on instruments and make up a nickname that they could remember in the following re-test administrations. In order to assess the test-retest reliability, researchers administered the translated version of EAQ to participants 20 days after its first administration. They took approximately 30 minutes to complete.

Adaptation Process

The Emotional Abuse Questionnaire was translated from English to Turkish with forward and backward translation methods (Hambleton, 2005, p. 12). First, the questionnaire was forward translated from English to Turkish by the first author, whose native language is Turkish and who is fluent in English. The translation was then reverse translated to the original English version by the second author, who is also native in Turkish and fluent in English. In the process of survey and questionnaire translations and adaptations, it is strongly suggested that translators be proficient in both languages and cultures involved, especially the target culture (Hambleton & Kanjee, 1995). It is also recommended that translators have knowledge of test construction as well as the subject matter being measured by the survey or questionnaire (Hambleton & Kanjee, 1995). As Turkish scholars who have lived in the United States for a long period of time, the authors meet these criteria. The first two authors, a bilingual marriage and family therapist and a counselor, were good fit for translations.

The reverse translated items and original items were then compared and corrected for discrepancies in vocabulary, phrasing, and semantics. Through consensus, minor changes were made in wording. When concerns or discrepancies arose, these items were discussed by the authors in terms of their cultural suitability. Once a consensus was reached, the appropriate changes were implemented. When the translation was complete, expert opinions were sought from Turkish and English language art and literature professors for the original items and on forward and reverse translations. These experts confirmed the final items as comparable to the original items.

Results

All statistical analyses were performed using the SPSS version 16 for Windows software package. As suggested by Jacobson and Gottman (1998), a principle-component analysis was performed to test the factor structure of each Emotional Abuse Questionnaire subscale.

Principal component analysis. Examination of scree plot and initial analysis as well as communality scores for the isolation subscale indicated that a one-factor solution was appropriate. Consequently, the solution was forced to a single factor, which yielded one principal component that represents isolation. Thirty two percent of the variance was explained by degradation component (eigenvalue=7.72). This subscale has 24 items, all of which loaded on to this component.

Table 1.

Summary of Principal Component Analysis of the Emotional Abuse Questionnaire Scale (N=95)

<i>Isolation</i>		<i>Degradation</i>		<i>Sexual abuse</i>		<i>Property damage</i>	
Item number	Factor loading	Item number	Factor loading	Item number	Factor loading	Item number	Factor loading
3	.74	37	.87	56	.87	63	.95
4	.72	41	.84	57	.77	61	.94
6	.64	36	.84	59	.67	62	.93
11	.61	49	.78	54	.58	66	.89
7	.61	32	.78	58	.51	65	.80
18	.58	31	.75	55	.50	60	.55
15	.56	48	.75	53	.39	64	.35
9	.55	40	.72				
8	.54	42	.71				
10	.54	26	.71				
16	.53	38	.70				
17	.52	29	.69				
23	.50	35	.69				
13	.41	28	.66				
19	.36	30	.64				
5	.34	25	.64				
2	.33	39	.58				
20	.33	26	.57				
21	.29	45	.49				
1	.21	50	.48				
		44	.47				
		52	.44				
		43	.41				
		27	.40				
		47	.33				
		34	.30				
		33	.25				
		51	.17				

For the degradation subscale, initial analysis indicated that a one-factor solution was appropriate. All 28 items were loaded to the degradation factor. This factor explained 39 percent of the total variance (eigenvalue=10.86). For the sexual abuse subscale, initial analyses of scree plot, KMO statistics, and communality scores also indicated that a one-factor solution was appropriate. Results yielded a seven-item principal component representing sexual abuse. Forty percent of the variance was explained by this sexual abuse component (eigenvalue=2.80). Finally, an examination of the initial solutions for the property damage subscale also indicated that a one-factor solution was appropriate. The results yielded a seven-item principal component, representing property damage. Sixty-five percent of the variance was explained by this component (eigenvalue=4.52).

Reliability. Item analysis was conducted via inter-item and item-scale correlations. The alpha value for the isolation subscale was .89, for the degradation subscale was

.92, for the sexual abuse subscale was .73, and for the property damage subscale was .86. These results are consistent with the original English subscales in that the sexual abuse subscale has the lowest reliability in both the original version and Turkish version, while the degradation subscale has the highest reliability in both versions. Overall, the internal consistency reliabilities indicate the coherence of the subscales.

To test reproducibility, test-retest reliability analyses were used. Reproducibility was assessed by administering the scale to participants twice within 20 days. Pearson correlation coefficients were used to determine the temporal stability of the four EAQ subscales. Results indicated that the test-retest reliability for isolation was .85, for degradation was .88, for sexual abuse was .75, and for property damage was .70. The translated and adapted EAQ appears to have adequate reliability, with reliability scores comparable to the original English version.

Validity. Criteria validity was evaluated by testing the association between conceptually related variables. It was hypothesized that the EAQ would negatively correlate with the DAS. As expected, EAQ total scores were negatively associated with DAS total scores ($r(93) = -.45, p < .001$). A detailed analysis of validity of the EAQ subscales based on correlations with the DAS subscales is presented in Table 2. As seen in the table, isolation was negatively correlated with relationship cohesion ($r(92) = -.22, p < .05$) and consensus ($r(94) = -.44, p < .05$). Similarly, degradation was negatively correlated with relationship cohesion ($r(93) = -.30, p < .05$) and relationship consensus ($r(93) = -.53, p < .001$). Property damage, on the other hand, was found to be negatively associated with relationship consensus ($r(93) = -.20, p < .05$). No association was found between property damage and cohesion. Sexual abuse was not correlated with any of the DAS subscales.

Table 2.

Pearson Product-Moment Correlations among Emotional Abuse Outcome Variables, Social Desirability, and Relationship Quality Variables

	Social desirability	Relationship consensus	Relationship cohesiveness	Dyadic adjustment
Isolation	-.13	-.44**	-.22*	
Degradation	-.03	-.53**	-.30*	
Sexual abuse	-.06	-.18	-.12	
Property damage	-.00	-.20*	.04	
EAQ total				-.45**

Conclusions and Recommendations

The purpose of this study was to adapt the Emotional Abuse Questionnaire (Jacobson & Gottman, 1998) into Turkish and to examine its psychometric properties in the Turkish cultural context. The construct of the Turkish Emotional Abuse Questionnaire was explored via its factor structure, and its criteria validity was tested by investigating the correlations between the Emotional Abuse Questionnaire and the Dyadic Adjustment Scale. Results indicated that the Turkish Emotional

Abuse Questionnaire is psychometrically comparable to the original English version of the scale. The reliability and validity analyses indicated that the Turkish version of the Emotional Abuse Questionnaire has acceptable psychometric qualities. A temporal stability was investigated by administering the Emotional Abuse Questionnaire twice. Furthermore, the Emotional Abuse Questionnaire subscales of isolation, degradation, sexual abuse, and property damage were found to measure relatively distinct constructs. The results further indicated that criteria validity was satisfactory; significant negative correlations were found between the Emotional Abuse Questionnaire and the Dyadic Adjustment Scale. These findings are consistent with previous studies conducted on a U.S. sample (Babcock, Jacobson, Gottman, & Yerington, 2000).

When we examined the results of the sexual abuse subscale of the Emotional Abuse Questionnaire, despite its acceptable reliability and factor structure, it yielded the lowest validity scores. This might be because of the differences of the cultural norms that regulate sexual activity. Predominant Turkish culture accepts sexuality within the institution of marriage. After the sexual revolution in the 1960s in Western countries, it was predicted that this revolution would spread around the world. However, there are cultural barriers in Turkey that prevent changing people's minds about sexuality (Kayır, 1995). For example, unlike many Western cultures, premarital sex is not common in Turkey. Yet, increasing permissiveness in premarital sexual activity in Turkey was observed, especially among women in large cities and university campuses (Öngen, 2006). Furthermore, education, economic independence, and living in large cities provide more opportunities and notions that encourage permissiveness (Werner-Wilson, 1998). In relatively underdeveloped parts of Turkey, many marriages are still commonly arranged by families, and sometimes women do not even see their husbands until their wedding night (Kayır, 1995). Furthermore, a premarital sexual relationship is a strong taboo in these areas. It is not uncommon to read news about honor-killing of women by their families. On the other hand, men tend to have more sexual freedom. In many cases, they are allowed to engage in sexual activity with other women and they are encouraged by society to be sexually active (Sakallı, Karakurt, & Uğurlu, 2001). This double standard against women reflects itself in the preference among most Turkish men for the chastity of the woman they marry (Yeni Yüzyıl, 1998). To this end, it is possible that premarital sexual attitudes played a role in the validity of the sexual abuse subscale of the Emotional Abuse Questionnaire. Given the conservative norms regarding premarital sex in Turkish culture, it is possible that respondents were unable or unwilling to respond in a meaningful way to the sexual abuse items. Since they may not have previous sexual experience, it may not be possible to respond to questions like "My partner is not sensitive to me during sex," "My partner pressures me to have sex after an argument," and "My partner intentionally hurts me during sex." Further research is needed to establish the validity of the sexual abuse subscale with married or sexually active couples.

On an issue such as abuse in dating relationships, it is possible that participants may select socially desirable responses. To account for this, social desirability bias was controlled by administering the short-version of the Marlowe-Crown Social

Desirability Scale. No significant correlation was found between social desirability and the scores on the Emotional Abuse Questionnaire, indicating that the responses were free from social desirability bias.

The interpretation of the results of this study must be viewed in light of its limitations. The Turkish Emotional Abuse Questionnaire is psychometrically sound and promising. However, in this study, a convenience sample of university students was used. Therefore, further research is needed to evaluate the qualities of the Turkish version with more diverse populations such as married adults and adults in dating relationships, the elderly, and sexually active adolescents. To further demonstrate the semantic equivalence of the Emotional Abuse Questionnaire, future work could focus on the dual language split-half (DLSH) method. Future research avenues also could include the association between mental health issues such as depression, anxiety and trauma, and emotional abuse in intimate relationships. Clinicians working with individuals, couples, and families also can use this scale to assess emotional abuse in client relationships. To conclude, the findings for the Turkish version of the Emotional Abuse Questionnaire show that it has sufficiently high reliability and good criteria validity, indicating acceptable psychometric qualities which are comparable to the original version.

References

- Akın, L., & Özaydın, N. (2005). The relationship between males' attitudes to partner violence and use of contraceptive methods in Turkey. *European Journal of Contraception and Reproductive Health Care, 10*, 199-206.
- Atalay, B., Konaş, M., Beyazıt, S., & Mağdenoğlu, K. (1992). *Türk aile yapısı araştırması (Investigation of Turkish family structure)* Devlet Planlama teşkilatı Sosyal planlama Genel Müdürlüğü Araştırma Dairesi, Yayın no: DPT: 313-SPGM: 412. Ankara: DPT matbaası.
- Babcock, J. C., Jacobson, N. S., Gottman, J. M., & Yerington, T. P. (2000). Attachment, emotional regulation, and the function of marital violence: Differences between secure, preoccupied, and dismissing violent and nonviolent husbands. *Journal of Family Violence, 15*, 391-409.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of Social Desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.
- Çelik, S. S., & Bayraktar, N. (2004). A study of nursing student abuse in Turkey. *Journal of Nursing Education, 43*, 330-337.
- Fışıloğlu, H., & Demir, A. (2000). Brief report: Applicability of the dyadic adjustment scale for measurement quality with Turkish couples. *European Journal of Consumer Research, 27*, 157-178.
- Follingstad, D., Rutledge, L., Berg, B., Hause, E., & Polek, D. (1990). The role of emotional abuse in physically abusive relationships. *Journal of Family Violence, 5*, 107-120.

- Fraboni, M., & Cooper, D. (1989). Further validation of three short forms of the Marlowe-Crowne scale of social desirability. *Psychological Reports, 65*, 595-600.
- Garcia-Moreno C, Jansen H. A., Ellsberg M., Heise L., & Watts H. C. (2006). Multicountry study on women's health and domestic violence against women study team prevalence of intimate partner violence: Findings from the WHO multi-country study on women's health and domestic violence. *Lancet, 368*, 1260-69.
- Gondolf, E. W., Heckert, D. A., & Kimmel, C. M. (2002). Nonphysical abuse among batterer program participants. *Journal of Family Violence, 17*, 293-314.
- Goodman, L. A., Koss, M. P., & Felipe-Russo, N. (1993). Violence against women: Physical and mental health effects. Part I: Research findings. *Applied and Preventative Psychology, 2*, 79-89.
- Hambleton, R. K (2005) Issues, Designs, and Technical Guidelines for Adapting Tests Into Multiple Languages and Cultures. In R.K. Hambleton, P. F. Merenda, C.D. Spielberger (Eds), *Adapting Educational and Psychological Tests for Cross- Cultural Assessment* (1st edition, pp. 3-38). Mahwah, NJ, Lawrence Erlbaum Associates.
- Hambleton, R. K. & Bollwark, J. (1991). Adapting tests for use in different cultures: Technical issues and methods. *Bulletin of the International Testing Commission, 18*, 3-32
- Hambleton, R. K., & Kanjee, A. (1995). Increasing the validity of cross-cultural assessments: Use of improved methods for test adaptations. *European Journal of Psychological Assessment, 11*, 147-157.
- Hamby, S. L. (1996). The dominance scale: Preliminary psychometric properties. *Violence and Victims, 11*, 199-212.
- Jacobson, N., & Gottman, J. (1998). *When men batter women: New insights into ending abusive relationships*. New York: Simon and Schuster.
- Joiner, W. (2008, October). Invisible violence, *Redbook, 211*, 190-196.
- Kağıtçıbaşı, C. (2000). *Cultural psychology*. İstanbul: Evrim Publication.
- Kağıtçıbaşı, C. (2003). Autonomy and relatedness in cultural context: Implications for self and human development. *Human Development, 46*, 145-150.
- Katz, J., & Arias, I. (1999). Psychological abuse and depressive symptoms in dating women: Do different types of abuse have differential effects? *Journal of Family Violence, 14*, 281-295.
- Kayır, A. (1995) Women and their sexual problems in Turkey. In S. Tekelli (Ed.), (pp. 288-305) *Women in modern Turkish society*. London: Zed Books Ltd.
- Leonard, K. E., & Senchak, M. (1996). Prospective prediction of husband marital aggression within newlywed couples. *Journal of Abnormal Psychology, 105*, 368-380.
- Marshall, L. L. (1999). Effects of men's subtle and overt psychological abuse on low income women. *Violence and Victims, 14*, 69-87.

- Murphy, C. M., & Hoover, S. A. (1999). Measuring emotional abuse in dating relationships as a multifactorial construct. *Violence and Victims, 14*, 39-53.
- O'Leary, K. D. (1999). Psychological abuse: A variable deserving critical attention in domestic violence. *Violence and Victims, 14*, 1-21.
- Öngen, D. (2006). Attitudes towards women: A study of gender and academic domain differences in a sample of Turkish university students. *Social Behavior and Personality: An International Journal, 34*, 467-486.
- Pejic, A., (2005). Verbal abuse: A problem for pediatric nurses. *Pediatric Nursing, 31*, 271-279.
- Romero, M. (1985). A comparison between strategies used in prisoners of war and battered wives. *Sex Roles, 13*, 537-547.
- Sakallı, N., Karakurt, G., & Uğurlu, O. (2001). Evlilik öncesi yaşanan cinsel ilişkiye ve kadınların evlilik öncesi cinsel ilişkide bulunmasına karşı tutumlar [Attitudes toward premarital sexuality and double standards against women, in Turkish]. *Tecrübi Psikoloji Çalışmaları, 22*, 15-29.
- Schumacher, J. A., & Leonard, K. E. (2005). Husbands' and wives' marital adjustment, verbal aggression, and physical aggression in early marriage. *Journal of Consulting and Clinical Psychology, 73*, 28-37.
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38*, 15-28.
- Stets, J. E. (1990). Verbal and physical aggression in marriage. *Journal of Marriage and the Family, 52*, 501-514.
- Straus, M. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics Scale (CTS). *Journal of Marriage and the Family, 41*, 75-88.
- State Institute of Statistics (SIS). (1990). *Population of Turkey, 1923-1994: Demographic Structure and its development*. Ankara: Turkey.
- State Institute of Statistics (SIS). (2008). *2007 census, population living in cities*. Ankara: Turkey.
- Sunar, D., & Fişek, G. (2005). Contemporary Turkish families. In U. Gielen & J. Roopnarine (Eds.), *Families in global perspective* (pp. 169-183). Boston: Allyn and Bacon./Pearson.
- Tolman, R. M. (1989). The development of a measure of psychological maltreatment of women by their male partners. *Violence and Victims, 4*, 159-178.
- UNECE. (2001). *Thematic atlas of Europe and North America*. Retrieved December 10, 2008, from <http://www.unece.org/stats/trends2005/employment.htm>
- Ural, T., & Özbirecikli, M. (2006). Is ethical judgment influenced by social desirability in responding? An analyze on Turkish accountants. *Sosyal Bilimler Enstitüsü Dergisi, 15*, 393-410.

- Wilson, M., Johnson, H., & Daly, M. (1995). Lethal and nonlethal violence against wives. *Canadian Journal of Criminology*, 37, 331-361.
- Werner-Wilson, R. J. (1998). Gender differences in adolescent sexual attitudes: The influence of individual and family factors. *Adolescence*, 33, 519-531.x
- Wijma, B., Schei, B., Swahnberg, K., Hilden, M., Offendal, K., Pikarinen, U. et al, (2003). Emotional, physical, and sexual abuse in patients, visiting gynecology clinics: A Nordic cross-sectional study. *The Lancet*, 361, 2107-2113.
- Yeni, Yüzyil (1998). *Türkiyede cinsellik*. Yeni Yüzyil Kitaplığı, p.34.
- Zook, A., & Sipp, G. J. (1985). Cross-validation of a short form of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 41, 236-238.

Duygusal Taciz Ölçeğinin Türkçe'ye Uyarlanması ve Psikometrik Özellikleri

(Özet)

Problem Durumu: Romantik ilişkilerde taciz evrensel olarak gözlemlenen bir sorundur. İlişkilerde taciz, fiziksel zarar verme şeklinde olabileceği gibi sözlü taciz, üstünlük kurmaya çalışmak, kontrol altına almak, alay etmek, aşağılamak ya da fiziksel zarar vermekle tehdit etmek şeklinde de kendini gösterebilir. Duygusal taciz, farklı araştırmacılar tarafından çeşitli şekillerde tanımlansa da, bazı ortak yönler belirlenmiştir. Genel olarak duygusal taciz, karşı tarafta duygusal hasar yaratma amacıyla zorlayıcı, caydırıcı davranışlar ve baskı uygulamak olarak tanımlanır. Duygusal taciz konusu son dönemlerde çalışılmaya başlanmış; ilişkilere ve bireylere verdiği zararlar açısından incelenmiştir. Duygusal taciz, psikolojik taciz, sözel taciz, psikolojik eziyet ve psikolojik saldırı gibi kavramlarla da betimlenmiştir. Duygusal taciz ciddi bir sorundur ve tacize uğrayan bireylere gerek fiziksel gerekse psikolojik olarak zarar vermektedir. Bu bağlamda, ilişki sorunlarının önemli bir göstergesi olan duygusal tacizin yapısı ve özelliklerinin Türk kültüründe ortaya konması gerekmektedir.

Araştırmanın Amacı: Bu çalışmanın amacı ilişkilerde duygusal taciz boyutlarını ve düzeylerini belirlemek üzere Duygusal Taciz Ölçeğini Türkçe'ye uyarlamak ve ölçeğin psikometrik özelliklerini Türk kültüründe test etmektir.

Araştırmanın Yöntemi: Duygusal Taciz Ölçeğinin geçerlik ve güvenilirlik çalışmalarına 82 kadın (%86) ve 12 (%14) erkek olmak üzere toplam 95 üniversite öğrencisi katılmıştır. Katılımcıların en genci 18, en yaşlısı 28 olmak üzere yaş ortalaması 20.58'dir ($Ss=1.94$). Ölçeğin Türkçe'ye çevirisinde ters çeviri yöntemi kullanılmıştır. Özgün maddeler öncelikle ilk yazar tarafından Türkçe'ye çevrilmiş; daha sonra bu Türkçe maddeler ikinci yazar tarafından yeniden İngilizceye çevrilmiştir. Daha sonra, bu

ikinci çeviri özgün maddelerle karşılaştırılmıştır. Buna ek olarak, Türkçe maddeler cümle yapısı, anlam ve kültürel uygunluğu konusunda incelenmiş ve fikir birliğine varılmıştır. Ayrıca, Türkçe maddeler, İngilizce ve Türk Dili öğretim üyelerine gösterilerek ölçeğin anlaşılabilirliği açısından görüşleri alınmış ve gerekli düzeltmeler yapılmıştır. Duygusal Taciz Ölçeğinin benzer ölçekler geçerliğini incelemek amacıyla Çift Uyum Ölçeği (Fışiloğlu ve Demir, 2000) kullanılmıştır. Varılan sonuçların sosyal beğenilirlik kaygısından etkilenmediğini doğrulamak amacıyla Marlowe - Crowne Sosyal Beğenirlik Ölçeği (Ural ve Özbircikli, 2006) kullanılmıştır. Bu ölçekler, araştırmaya katılmaya gönüllü üniversite öğrencilerine uygulanmıştır. Veriler öğrencilerden doğal sınıf ortamında, araştırmacı tarafından tek oturumluk uygulamalarla toplanmıştır. Araştırmaya katılan öğrencilere araştırmanın amacı ve yapılan uygulamayla ilgili bilgi verilmiştir. Veri toplama işlemi bitirildikten sonra elde edilen veriler uygun istatistiksel işlemlerin yapılabilmesi için hazır hale getirilmiştir. Duygusal Taciz Ölçeği geçerlik çalışmaları kapsamında, yapı geçerliği faktör analizi ile incelenmiştir. Benzer ölçekler geçerliğinde, Duygusal Taciz ile Çift Uyum Ölçekleri'nden alınan toplam puanlar arasındaki ilişki Pearson Korelasyon Katsayısı hesaplanarak bulunmuştur. Duygusal Taciz Ölçeği'nin değerlendirilmesi kapsamında, Cronbach-alfa katsayısı ve madde toplam korelasyonları hesaplanmıştır. Ayrıca ölçeğin kararlılık katsayısını bulmak amacıyla ölçek aynı katılımcı grubuna 20 gün içerisinde iki defa uygulanmış, ve bu iki uygulamada alınan sonuçlar Pearson Korelasyon Katsayısı kullanılarak karşılaştırılmıştır.

Araştırmanın Bulguları: Ölçeğin yapı geçerliğini belirlemede, aynı yapıyı ya da niteliği ölçen değişkenleri bir araya toplayarak ölçülen değişkenleri az sayıda faktör ile açıklamayı amaçlayan temel bileşenler analizi uygulanmıştır. Ölçeğin geçerlik çalışmasında benzer ölçekler geçerliği de uygulanmıştır. Çift Uyum Ölçeği kullanılarak gerçekleştirilen ölçüt geçerliği çalışmasında, Pearson korelasyon katsayısı .45 ($p < .001$) olarak bulunmuştur. Bu sonuç Duygusal Taciz Ölçeğinin geçerli sayılabilmesi için yeterli olduğunu göstermektedir. Ölçeğin güvenilirliği, Cronbach-alfa katsayısı ile değerlendirilmiştir. Buna göre Duygusal Taciz Ölçeğinin alt ölçeklerinin güvenilirlik katsayısı, tecrit etme için .89, aşağılama için .92, cinsel taciz için .73 ve mala zarar verme için .86 olarak bulunmuştur. Ölçeğin kararlılık katsayısını bulmak amacıyla test-tekrar test yöntemi uygulanmıştır. Araştırma örneklemine ölçek yirmi gün arayla iki kez uygulanmış ve Pearson korelasyon katsayısı altölçeklerde aşağılama için .88, tecrit etme için .85, cinsel taciz için .75 ve mala zarar verme için .70 olarak bulunmuştur. Duygusal Taciz Ölçeğinin güvenilirlik katsayıları ölçeğin, üniversite öğrencilerinin romantik ilişkilerinde duygusal taciz düzeylerini ölçmek amacıyla güvenle kullanılabileceğini göstermektedir.

Araştırmanın Sonuçları ve Önerileri: Duygusal Taciz Ölçeğinin geçerlik ve güvenilirlik çalışmalarından elde edilen bulgular, Duygusal Taciz Ölçeğinin üniversite öğrencilerinin romantik ilişkilerinde duygusal düzeylerini geçerli ve güvenilir bir şekilde ölçmek amacıyla kullanılabileceğini

göstermektedir. Duygusal Taciz Ölçeğinin geçerliği, faktör analizi ve benzer ölçekler geçerliği ile belirlenmiştir. Temel bileşen analizi sonucunda ölçeğin çok boyutlu olduğu görülmüştür. Ancak, cinsel taciz alt ölçeği için geçerlik katsayısı ve faktör yapısı bulgularının kabul edilebilir düzeyde olmalarına rağmen, bu değerlerin diğer altölçeklere kıyasla daha düşük güvenilirlik gösterdiği gözlenmiştir. Bu durum, Türk kültüründe evlilik dışı cinsel ilişkilerin hala tabu olarak görülmesi ile açıklanabilir. Genel olarak, Duygusal Taciz Ölçeğinin güvenilirlik çalışmaları sonuçlarına göre, Cronbach-alfa, madde toplam korelasyonu katsayıları ile test-tekrar test yöntemiyle elde edilen korelasyon katsayılarının kabul edilebilir düzeyde olduğu görülmektedir. Araştırma sonuçlarının sosyal beğenilirlik kaygısından bağımsız olduğunu test etmek amacıyla, Marlowe- Crown Sosyal Beğenilirlik Ölçeği kullanılmıştır. Duygusal Taciz Ölçeği ile Marlowe- Crown Sosyal Beğenilirlik Ölçeği arasında istatistiksel olarak anlamlı bir ilişki bulunmamıştır. Bu sonuç, maddelerin sosyal beğenilirlikten bağımsız olduğunu göstermektedir.

Anahtar Sözcükler: Duygusal taciz, Duygusal Taciz Ölçeği, Çift Uyum Ölçeği, romantik ilişkiler.

Counselors and Administrators: The Collaborative Alliance in Three Countries

Fidan Korkut-Owen*

Dean W. Owen**

Victor Ballesterero***

Suggested Citation:

Korkut-Owen, F., Owen, D., & Ballesterero, V. (2009). Counselors and administrators: the collaborative alliance in three countries . *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 23-38.

Abstract

Throughout most of the past 100 years there have been major paradigm shifts in the structure and organization of schools. Perhaps one of the most widespread changes has been the integration of counseling into public schools. This concept, originally defined as "guidance," evolved first in the United States but has been widely adopted throughout the world in industrialized countries. As the concept of education has been more widely defined, schools have sought not only to teach information but also to prepare students for a productive future and meaningful participation in the adult world. This has led to the adoption of school counseling programs as an integral component of the public education system. The school counselor has thus joined the classroom teacher and school administrator as the third major component of the professional educational staff in many nations across the Europe, Asia, Australia, and the Americas. The role and function of the school counselor have evolved in quite different ways in different countries and this article explores this evolution as well as the current status of school counseling as it exists in Costa Rica, Turkey, and the United States. These three countries represent the countries of origin of the three authors.

In spite of great historical, cultural, and political differences, each of the three comparison countries has integrated school counseling into its respective educational system and has done so in unique ways that express much about the cultural, social, and political expectations of that country. Growing from the traditional model of schools, composed of teachers and principals, has come a third force which includes counselors and other specialized, supportive services and staff that only a few generations ago were regarded as a frivolous luxury. By examining the history, evolution,

* Professor, Hacettepe University, Faculty of Education, korkut@hacettepe.edu.tr

** Professor, Morehead State University, College of Education, d.owen@moreheadstate.edu

*** Associate Professor, Morehead State University, College of Education, v.ballesterero@moreheadstate.edu

and current role and function of the school counselor from the vantage point of three quite different countries, it is possible to not only recognize differences and similarities but also to observe the extent to which counseling has become fully integrated into public education. A school counselor, in each of the three comparison countries, is called upon to perform a job that is increasingly defined by government policy, interpreted by school administrators, and critically evaluated by parents, teachers, and school administrators. The United States, Turkey, and Costa Rica represent three distinct stages in the integration and evolution of counseling services within public education. Information about the role and function of the school counselor and how counselors and principals relate in promoting effective schools is reviewed and discussed for each of the three comparison countries.

Keywords: School counselor, principals, Costa Rica, Turkey, USA

School counseling has emerged over the last century to become a clearly defined profession within the educational structure of many industrialized countries. The profession of school counseling is thought to have first begun in the United States under the title of “guidance” and evolved out of a concern for poverty stricken youth (Gladding, 2004). Counseling was integrated into the public schools in the United States during the 1920s, and over the past century the concept of the school guidance counselor has spread and become integrated into the public schools of many other nations. The school counselor has joined teachers and principals to become a significant third component of the professional team serving the educational and developmental needs of school children. The variety of cultures, economic bases, religious values, and the availability of resources are only some of the reasons why school counseling has evolved in uniquely different ways around the world (Hatch & Bowers, 2005).

The countries to be examined in this analysis are the United States, Turkey, and Costa Rica and represent the countries of origin of the three authors, each of whom has extensive educational and vocational experience in at least two of the comparison countries. These three countries also represent distinct stages in the evolution of school counseling. In the United State counseling has been an integral component of the public education for nearly a century, whereas school counseling in Costa Rica is at a much earlier stage in its development. Our goal is to compare and contrast the role of the school counselor across these widely disparate countries with a particular emphasis on administrative structures and how the roles of the school counselor and the school principal are interrelated. The interrelated nature of both school counseling and school administration will be portrayed as an essential and defining quality of schools that effectively and efficiently facilitate learning among students.

As a profession, school counseling is perhaps most fully developed in the United States, while in Turkey and Costa Rica the profession may be regarded to be still evolving and seeking to more firmly establish itself within the structure of the educational system. This cross-national view of school counseling can only represent a snapshot of how these three educational systems have elected to integrate school

counseling into public education. This integration has been influenced by many factors and a comparison among the three countries provides insight into how the profession has found its place, form, and expression within the education community of three widely divergent countries. An analysis of how the counseling and administrative roles have interacted with, complimented, and in some cases conflicted with one another can provide clarity on the process of finding the optimal collaborative relationship for effective schools. An understanding of the roles of teachers and administrators is clearly established by tradition, but the role, function, and even the training of the school counselor remain poorly understood. In many cases it is not even well understood by other school professionals, particularly in countries that have only recently sought to integrate counseling within the public schools.

Overview of the comparison countries

By far the largest of the three countries is the United States, with an estimated population in excess of 300 million of which slightly more than 20% are younger than 14 years of age (United States Census Bureau, 2008). The literacy rate in the United States is currently 99%, with nearly 56 million students enrolled more than 87,619 elementary and secondary schools (United States Department of Education, Institute of Education Science, 2008). Compulsory education in the United States arose with the development of its formal educational system in the mid 19th century and, although the duration of required education varies from state to state, the majority require attendance until the age of 16 years (Johnson, Collins, Dupuis, & Johansen, 1985).

The current organization of the United States public education is influenced by a complex mixture of local, state, and national funding. At the national level, the U.S. Department of Education has the primary responsibility for determining policies and regulations for federally funded initiatives that state and local school boards must adopt in order to enjoy federal funding. Each of the 50 states similarly has a department of education which regulates the educational system in their respective states. This includes primary funding, setting training standards for teachers and other school professionals, adopting a state-approved curriculum, and a variety of other matters. Finally, within each state there are organizational units known as school districts, each of which has an elected school board that sets local policies and regulations for the schools within its control (United States Department of Education, 2008).

Counseling in the United States began nearly a century ago and has since evolved into a separate and distinct profession apart from other types and forms of psychological services (Owen, 2007). Counseling traces its roots to social concerns arising from the crowded conditions faced by the poor and immigrant populations in the large eastern cities of New York and Boston in the late 1800s. Frank Parsons' pioneering work in 1909, with the creation of the Boston Vocational Bureau and the publication of his book *Choosing a Vocation*, effectively began the guidance movement in the United States. The school counseling profession traces its origin to Jesse B. Davis, who is considered as the first to provide a systematic school-guidance

programme. In 1907, he became the principal of a high school and encouraged the school's English teachers to use compositions and lessons as a vehicle to teach career information, encourage career exploration, develop character, and avoid behavioral problems (Gladding, 2004).

Over the next 80 years counseling increasingly became integrated into the basic fabric of the American educational system. Continued professional development occurred throughout the last half of the 20th century as national professional associations evolved to further define and promote the training standards, role, and function of the school counselor. Led by Campbell and Dahir (1997), who consulted with school counselors at all levels, the American School Counselor Association (ASCA) created the ASCA National Standards for School Counseling with three core domains (Academic, Career, Personal/Social), nine standards, and specific competencies and indicators for K-12 students (Hatch, & Bowers, 2005).

School-counselor training in American universities typically involves the completion of a master's degree with coursework designed to prepare a candidate to perform the role of school counselor as outlined by the departments of education of each state. Although all state departments of education have adopted essentially the same standards, small differences exist among states. Each state examines and accredits counselor-education programs at universities within the state. When candidates complete the graduation requirements and make appropriate applications, a certificate or license is granted which permits them for employment in the public schools of that state. Currently, most states in the United States require that school counselors must first be trained and experienced as a classroom teacher before beginning training as a school counselor.

The training programs for U.S. school counselors vary according to the individual state requirements, but typically involve a minimum of 30-48 semester hours of study in an approved counselor education program. At the present time there are 523 colleges and universities offering programs of education in counselor preparation and 139 offering doctoral training (Peterson's Graduate Planner, 2007). Candidates completing programs may elect to be employed at elementary schools, middle schools, and high schools depending upon their certification by the state department of education.

The specific role expectations outlined for U.S. school counselors also vary by state, but the vast majority of states have adopted the ASCA National Model, a standard framework for school counseling programs. This model represents a comprehensive statement of expectations and roles, and defines the school-counseling program in terms of student development throughout their employment in public schools. This developmental approach to the delivery of student-counseling services addresses the three primary areas of concern including academic development, career development, and personal/social development (ASCA, 2008).

The ASCA national standards for school-counseling similarly defines a wide range of inappropriate duties for U.S. counselors. This includes many tasks that would be considered routine clerical functions, unskilled monitoring functions such

as supervising student study periods, and acting as a substitute teacher. This set of national standards has done much to define the professional role and function of the of the school counselor in the United States (ASCA, 2008). As the result of a lengthy period of development and evolution, the profession and practice of school counseling is fully mature and deeply integrated into U.S. public education.

The country of Turkey has a current population estimated to be in excess of 70.5 million. Turkey's population is also remarkably young, with more than 26.4 % being 14 years of age or younger (Turkish Statistical Institute, 2007). The literacy rate for individuals aged 15 and older is currently estimated to be in excess of 87% (United Nations Development Programme, 2008). Public school enrollment during the 2007-2008 academic year was in excess of 12 million students attending nearly 35,000 primary and secondary schools (Turkish Statistical Institute, 2008). The current compulsory education laws in Turkey require that students attend public school through the 8th grade.

In Turkey, all formal and non-formal education is implemented under the supervision and control of the state (Milli Eğitim Bakanlığı, MEB) through the Ministry of National Education (MONE). MONE has the primary responsibility for all facets of the Turkish public education system including staffing, curriculum, and oversight of both elementary and secondary programs. There are MONE representatives in all 81 of the officially recognized municipalities across the country.

The current Turkish counseling movement and its programs have been largely influenced by American counseling programs, models, and philosophies. Between 1950 and 1956, the visits of some American counselor educators and the efforts of Turkish pioneer counselor educators trained in the United States played a significant role in bringing counseling concepts into Turkish education. It isn't until the period after 1970, however, that professional counseling really began to be practiced in schools. In 1982 the Turkish universities began admitting students to four-year bachelor of education programs with a major in guidance and counseling (Doğan, 1999).

Currently in Turkey a total of 66 counseling programs are housed in 39 universities. Only two private universities offer counseling programs. There are 31 undergraduate counselor education programs (school counselors are trained at the undergraduate level), 21 master's programs, and 14 doctoral degree programs (Çınarbaş, Korkut-Owen and Çiftçi, in press). Candidates completing initial training at the bachelor's degree level are required to complete an exam called Kamu Personel Seçme Sınavı -Teacher Version (Civil Service Personel Selection Exam): a nationally administered and highly centralized exam. Candidates who successfully complete this exam are deemed qualified to work as school counselors, although in Turkey "Guidance Teacher" is the term used by the Ministry of Education. Successful candidates may be assigned to vacant counseling positions throughout the country or they may apply to a private school or a university preparatory school (MONE, 2008).

Within the Turkish educational system the expectations for counseling programs and the counselors are broadly defined by MONE. The expectations seek to provide

guidance and psychological counseling as a systematic assistance process provided for groups or individuals in order to ensure that guidance teachers meet the students' needs and solve their problems (OECD-The Organization for Economic Cooperation and Development, 2005). According to MONE (2001), counselors' roles and functions are defined as providing consultation to classroom teachers in administering school guidance and counseling programs; providing necessary resources, materials, and tools for measurement; evaluating programs together with other teachers and improving programs as needed; informing parents about the programs; and designing guidance and counseling programs based on school's needs and conditions. The original regulation that defined counselors' roles and functions (MONE, 2001) contained specific statements describing specific job functions that were inappropriate for school counselors. This regulation prohibited school counselors from performing clerical functions, taking on broad administrative responsibilities, performing as a substitute teacher in the absence of normal classroom teachers, proctoring examinations, and performing "hall monitor" duties.

After 2001, MONE implemented developmental and preventive ideas into counseling programs by regulation. During the 2000-2001 academic year, a national guidance program was pilot tested for grades 9-11, and this concept was further extended in 2006 with the adoption of the Elementary and Secondary Schools Classroom Guidance Program (İlköğretim ve Ortaöğretim Kurumları Sınıf Rehberlik Programları) (MONE, 2006). This nationally adopted scheme for school-guidance services focused on educational, career, and personal/social counseling and guidance, and was extended to include grade 12 for the 2008-2009 academic year.

The Central American country of Costa Rica is the smallest of the three comparison countries and has a population estimated to be slightly in excess of 4 million, as of July 2008. This population is, again, remarkably young, with more than 27% reported to be under the age of 14 years and only about 6% over the age of 65 years (Central Intelligence Agency, 2008). Costa Rica has one of the highest literacy rates in all of Central America at 96% (Petersen, 2008) and during 2008 there were more than 3,670 primary and secondary schools educating more than 919,400 public school students (Ministry of Public Education, 2008). Compulsory school attendance in Costa Rica requires that students attend through the 9th grade, although this is rarely enforced.

The Costa Rican public education system is a highly centralized national system administered by the Ministry of Public Education- MPE (Ministerio De Educacion Pública). Costa Rica's central government manages all of the public educational institutions and controls and approves the curriculum for both formal and non-formal education (Petersen, 2008). The seven provinces in Costa Rica each have an administrator and school boards (United States University Directory, 2008).

The origins of counseling as a component of Costa Rican education can be traced to the early 1950s; counseling did not appear as a specialization until the University of Costa Rica offered the first program of counselor education in 1956. The country's National Ministry of Education Department of Counseling was established in 1957. Despite the efforts over the past 50 years, counseling remains only partially available

to Costa Rican students in public schools where such services are offered only at the secondary level.

In Costa Rica, bachelor-level counselor-education programs are offered at 34 universities (2 public and 32 private). After completion of the bachelor's degree, students have the option of pursuing a graduate degree called *licenciatura* which takes two additional years of study and the completion of a dissertation. After this degree, students may complete their master's degree requiring an additional two years. There are currently more than 22 universities offering master's degrees in counseling (according to Costa Rican counselor educator Dr. Victoria Morales, personal communication, December, 2008). In Costa Rican public schools, counselors are employed at the secondary level only, although at private schools a second counselor is typically employed to work with elementary students (Alfaro, 2005).

Within the Costa Rican educational structure a law entitled "*Reglamento de Carrera Docente*" (Regulations of the Teaching Career) governs nearly all aspects of those professionals working in public and private educational settings, including teachers, principals, and counselors. This law was enacted in 1972, and defines the rights, responsibilities, and regulations for selection, tenure, and dismissal of school counselors (MPE, 2008). The regulation defines areas in which students are supposed to benefit as the result of a counseling program, but the specific role and function of the Costa Rican school counselor remains vaguely defined and allows great latitude for interpretation. The statements of role definition and the delineation of job functions for counselors fail to specify any specific inappropriate job functions.

The Counselor's Struggle for a Well-defined Role

Much of what has been presented thus far portrays school counselors as having much in common with one another regardless of the country in which they work. Counselors have become established as the third element of the school's professional staff by joining the teaching faculty and those charged with school administration. In each of the three comparison countries the counselor has been the most recent to join the professional staff.

The traditional roles and functions of classroom teachers are deeply rooted in hundreds of years of cultural traditions, as are the roles of administrators. There is an idiom used in English to describe a person or a situation that is not easily defined or classified. That idiom is "neither fish nor fowl," and this certainly describes the role of the school counselors who are neither teachers in the traditional sense nor administrators who provide oversight and direction of the school and its curriculum. Counselors are frequently regarded by teachers as "administrators" and by school administrators as substitute teachers, utility clerical workers, disciplinarians, test administrators, etc. Despite attempts to clearly define the role, scope, and function of the school counselor through law or regulation, there is evidence to suggest that the school counselor's position is difficult to manage. It is an ongoing struggle as school counselors attempt to juggle the often-conflicting expectations of teachers and administrators with the role for which they have prepared during their years of training.

In each of the three countries the school principal has the primary responsibility for oversight and school operations, and is traditionally the individual responsible for all school programs, including the school guidance/counseling program. Principals hold the most powerful position in schools because they typically identify, select, and appoint staff; determine school organizational and functional structure; and dictate the roles and functions of each school staff member (Taylor, 2002). The principal, then, is viewed in all three countries as the pivotal individual who determines the day-to-day activities of the school's professional staff. This relationship between school principal and school counselor will be more thoroughly examined.

Among the three countries it is Costa Rica which, perhaps, has the least well-defined role for the school counselor. The role and function of the school counselor in Costa Rican schools have yet to be investigated in a systematic fashion. In the public school system the counselor is largely regarded as a "vice" principal by the teaching faculty. Counselors are frequently assigned by the principal to serve as substitute teachers, perform test monitoring duties, and generally function as a roving, utility teacher. In Costa Rica, the school counselor has perhaps the greatest risk of being asked to perform non-counseling related job duties because the counselor's role is so vaguely defined and because the principal has traditionally been regarded as a nearly absolute authority figure whose decisions are rarely questioned (Ballestero, 1987).

The relationship between Turkish counselors and principals has only recently become a topic of research. Several researchers have reported that the roles and functions of school counselors are not clear, partially because administrators do not know counselors' roles and functions (Nazlı, 2007). Despite regulations defining the roles and functions of Turkish school counselors (MONE, 2001), counselors frequently assert that they are frustrated because of the competing demands and the expectations of the school's principal. In another study, teachers and administrators were found to have a confused and poorly defined idea of the role of the school counselor (Özaydın, 2001). Korkut-Owen and Owen's (2008) results suggested a remarkably similar pattern of role expectations and valued counselor activities among administrators and school counselors. An unexpected finding was that the administrators held a more favorable opinion toward the school-guidance program than did the school counselors. In the same study, administrators and counselors both overwhelmingly indicated that guidance programs were essential or highly important to the overall school mission and role within the community. While counselors and administrators appear to agree on the importance of a school guidance program and of school-counseling, there still seems to be confusion and frustration over the specific roles and functions which principals expect counselors to perform.

The recognition of the importance of counseling is illustrated by the extent to which counseling and guidance issues are embedded within Turkish elementary school branch curricula (Aşkar, Paykoç, Korkut, Olkun, Yangın, & Çakıroğlu, 2005). Söker (2007) reported that elementary school teachers believe that guidance services are useful for class management and for teaching students how to study. These studies seem to suggest that, although there is some confusion about roles and

functions of counselors and counseling services, counseling services are considered to be a valuable and important component of education. Since 1990, in an attempt to reduce the uncertainties about the role and functions of school counselors, the Turkish Ministry of National Education has provided in-service training to more than 450,000 teachers and administrators. This initiative was begun in order to more fully develop guidance and psychological counseling services in the Turkish education system and to help ensure that teachers and administrators gain the necessary understanding and competencies about school guidance services (OECD, 2005). Another indicator of the essential value of school counseling services is the fact that all students studying for degrees in education have been required, since 1997, to take a guidance course during their training (YÖK: Higher Education Council, 1997).

It is not surprising that the relationship between school counselors and school principals has been most thoroughly investigated in the United States, where counseling emerged much earlier. In a recent study by Jonson, Milltello, and Kosine (2008), using a Q-sort methodology, 39 professional counselors and principals were asked to sort 45 opinion statements related to the relationship between the two. Four factors emerged which described the essence of the relationship. These factors were: (1) Working Alliance, (2) Impediments to Alliance, (3) Shared Leadership, and (4) Purposeful Collaboration. With many states in the United States currently undergoing educational reform initiatives, the importance of collaborative leadership between the two positions and with the teaching faculty is seen as essential. These results were generally consistent with those reported by Ponec and Brock, (2000). In an attempt to document how future school principals view the roles of professional school counselors, studies by Chata and Loesch (2007) and Fitch, Newby, Ballesterio, and Marshall, (2001) effectively demonstrated that future school administrators have a relatively clear view of appropriate school-counselor responsibilities that is congruent with current counselor training.

A number of recent studies have sought to investigate the differences between administrators and counselors in regard to their views of the roles and functions of a school counselor. The results would seem to indicate that in U.S. schools, both groups share a remarkably similar belief about what and how counselors should provide services (Kirchner, & Setchfield, 2005; Owen, 2005; Zalaquett, 2005). Kaplan (1995), however, presented a far different viewpoint, describing multiple reasons for potential conflict between counselors and principals that result from the fact that both are trained using very different paradigms. He points to the fact that the focus of their efforts are quite different: principals adopt a business and systems viewpoint, reflecting the importance of decisions and actions for the school and community, whereas counselors tend to focus more on micro-level matters as they seek to facilitate change and development in individual students, particularly those experiencing difficulties.

Some research (Kaplan, 1995; Ponec, & Brock, 2000) seems to indicate that for a school counselor to be effective there must be demonstrable support for the school counselor and the school's guidance programs. It is because of this that Hatch and

Bowers (2005) strongly encourage principals and counselors to work collaboratively toward a common goal as defined by the ASCA national model. Two recent studies suggest that the relationship between the counselor and the principal is an important component of a successful and smoothly running school (Dollarhide, Smith, & Lemberger, 2007; Monteiro-Leitner, Asner-Self, Milde, Leitner, & Skelton, 2006).

Conclusion

This article has sought to compare three countries, each of which has adopted school counseling as a major component of its educational system. Each country, however, is an example of a very different stage in the professional development and maturity of the profession. Not surprisingly, the United States, where counseling arose nearly a century ago, has the most developed and mature counseling profession of the three. With well-developed programs of education, training and performance standards, national accreditation, and well-established scholarly associations to support the profession, school counseling can evolve to become an integral component of a country's educational system. The efficacy and importance of counseling in the United States' schools is well documented.

Turkey has, in recent years, made great strides in the integration of school counseling with the adoption of a national school-guidance program that is likely to continue to influence the profession and its training for the future. Having adopted a Western model of counseling services as far back as the 1950s, the counseling profession in Turkey is rapidly expanding and developing its own unique identity. With a rapidly growing professional association and an increasing interest in establishing national standards for training and performance, there is every reason to believe that school counseling in Turkey will continue to evolve into a stronger and more well-defined profession.

Costa Rica, the smallest of the comparison countries, continues to promote the importance of counseling as a significant component of its educational system. Although school counseling in Costa Rica has seen the least professional development so far, the profession continues to mature and to be integrated into its educational system, notwithstanding its relatively diminutive size.

Although counseling has become an integral component of the educational system in each of these countries, there is ample evidence to suggest that when the roles for teachers, counselors, and administrators are clearly delineated and understood the overall effectiveness of the school is improved. These roles are perhaps most clearly defined and understood in the United States, where counseling is viewed as an essential component in creating successful schools. School counseling in the United States has also benefitted from the strong advocacy of national professional associations that have done much to define the roles and functions of the school counselor. Turkey, understanding the importance of the school counselor, has recently taken substantial steps in clarifying both the role and function of counselors and is educating both administrators and teachers in those roles and functions. The organization and development of a national professional association for counseling in Turkey is continuing and is still seeking its voice as a

strong and powerful advocate for counseling in schools. In Costa Rica school counselors are mandated for only secondary level schools and have the least clearly defined roles. School counselors therefore continue to struggle to be regarded as more than utility and substitute teachers or test monitors.

Ultimately the structure, form, role, and function of school counseling have evolved differently in each of the comparison countries. These differences may be attributable to many factors including the basic structure of public education, from highly centralized national ministries in the case of Turkey and Costa Rica to educational systems that are highly decentralized with far more local control as in the United States. Long standing traditions, the presence of strong and vocal professional associations that advocate for school counseling and school counselors, and cultural values and social needs that become expressed through a nation's education system all have contributed to differences among the comparison countries. In spite of these differences, in each of these countries the profession of school counseling has found a place among the teachers and school administrators as the third major force in their respective national educational systems.

References

- Alfaro, J.M. (2005). *Autoevaluacion con fines de acreditacion y cultura de la calidad en la educacion superior en Costa Rica* [Self evaluation for accreditation purposes and culture of the quality of higher education in Costa Rica]. Retrieved December 11, 2008, from www.pmaet.gob.ni/Seminarios/Presentaciones/AdeAyCdelaCalidad%20Costa-Rica%20Dr%20Jorge%20Mora.pdf
- American School Counselor Association. (2008). *ASCA National Model*. Retrieved December 12, 2008, from www.schoolcounselor.org/files/appropriate.pdf
- Aşkar, P., Paykoç, F., Korkut, F., Olkun, S., Yangın, B., & Çakıroğlu, J. (2005). *Yeni öğretim programlarını inceleme ve değerlendirme raporu* [Review and evaluation report on new curriculum]. Unpublished report.
- Ballesteros, V. (1987). *Comparacion del Curriculum de Escuelas Publicas y Privadas en la Zona Norte*. [Comparison of the Curriculum of Private and Public Schools in the North Zone]. Unpublished master's thesis, Universidad de Costa Rica. Mayo.
- Chata C. C., & Loesch, L. C. (2007). Future school principals' views of the roles of professional school counselors. *Professional School Counseling, 11*(1), 35-41.
- Central Intelligence Agency. (2008). *The World Factbook: Costa Rica*. Retrieved December 7, 2008, from <https://www.cia.gov/library/publications/the-world-factbook/print/cs.html>
- Çınarbaş, D.C., Korkut-Owen, F., & Çiftçi, A. (In press). Counseling in Turkey, (Eds: Puncy Heppner, Stefania, Ægisdóttir, Alvin Leung, Kathryn Norsworthy, and Larry Gerstein). *Handbook of cross-cultural counseling: Cultural assumptions and practices worldwide*. Sage Publications.
- Doğan, S. (1999). Counseling in Turkey: Current Status and Future Challenges. *Education Policy Analysis Archives, 6*(12).

- Dollarhide, C. T., Smith, A. T., & Lemberger, M. E. (2007). Critical incidents in the development of supportive principals: Facilitating school counselor-principal relationships. *Professional School Counseling, 10*(4), 360-369.
- Fitch, T., Newby, E., Ballester, V., & Marshall, J. L. (2001). Counselor preparation: Future school administrator's perception of school counselor's role. *Counselor Education and Supervision, 41*, 89-99.
- Gladding, S. T. (2004). *Counseling: A comprehensive profession* (5th ed). Upper Saddle River, NJ, Pearson Prentice Hall.
- Hatch, T., & Bowers, J. (2005). *The ASCA National Model: A framework for school counseling programs*. Alexandria, VA, American School Counselor Association.
- Higher Education Council (1997). *Eğitim fakültelerinde uygulanacak yeni programlar hakkında açıklama* [Announcement of new programs for faculties of education]. Retrieved December 21, 2008, from www.yok.gov.tr/egitim/ogretmen/aciklama.doc.
- Johnson, J. A., Collins, H. W., Dupuis, V. L., & Johansen, J. H. (1985). *Introduction to the Foundations of American Education* (6th ed.). Boston: Allyn and Bacon.
- Johnson, C., Milltello, M., & Kosine, N. (2008). Four views of the professional school counselor-principal relationship: A Q-Methodology study. *Professional School Counseling, 11*(6), 353-361.
- Kaplan, L. S. (1995). Principals versus counselors: Resolving tensions between different practice models. *School Counselor, 42*(4), 161-268.
- Kirchner, G., & Setchfield, M. (2005). School counselors' and school principals' perceptions of the school counselor's role. *Education, 126* (1), 10-16.
- Korkut-Owen F., & Owen, D. W.(2008). Okul psikolojik danışmanlarının rol ve işlevleri: yöneticiler ve psikolojik danışmanların görüşleri [School Counselor's Role and Functions: School Administrators' and Counselors' Opinions] . *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 41*(1), 204-218.
- Ministry of National Education. (2001). *Rehberlik ve psikolojik danışma hizmetleri yönetmeliği* [Regulations for counseling and guidance services]. T.C. Resmi Gazete [Republic of Turkey Legislative Record] 17 Nisan 2001, 24376.
- Ministry of National Education. (2006). *İlköğretim ve orta öğretim kurumları sınıf rehberlik Programı* [Elementary and secondary schools classroom guidance program]. MONE, Special Education and Guidance Service General Directory. Retrieved November 12, 2008, from www.orgm.meb.gov.tr
- Ministry of National Education. (2008). *MEB Personel Genel Müdürlüğü 2008-2 Öğretmenlik İçin Başvuru Ve Atama Kılavuzu* [MONE Human Source Directory 2008-2, Guide for Application and Placement for Teacher Position]. Retrieved November 25, 2008, from www.personel.meb.gov.tr/upload/2008-2ogretmenlikinbasvuruveatamakilavuzu_temmuz2008.pdf
- Ministry of Public Education. (2008). *Reglamento de carrera docente* [Regulation for teaching career]. Retrieved December 11, 2008. from www.mep.go.cr
- Ministry of Public Education. (1998). *Manual del programa de orientacion educativa de tercer ciclo* [Regulations of the school counseling program of the third cycle], Departamento de Orientacion, San José, Costa Rica, EUNED.

- Monteiro-Leitner, J., Asner-Self, K. K., Milde, C., Leitner, D. W., & Skelton, D. (2006). The role of the rural school counselor: Counselor, counselor-in-training, and principal perceptions. *Professional School Counseling, 9*(3), 248-251.
- Nazlı, S. (2007). *Okul yöneticilerinin rehberlik ve psikolojik danışma hizmetlerini algılamaları* [School principals' perception of the guidance and counseling service]. *Eğitim Araştırmaları Dergisi, 26*, 281-290.
- Organization for Economic Co-operation and Development. (2005). *Reviews of national policies for education: basic education in Turkey background report*. Retrieved November 12, 2008, from www.oecd.org/edu/reviews/nationalpolicies
- Owen, D.W. (2005). School administrators' perceptions of counselor roles and functions. *Journal of Kentucky Counseling Association Journal, 24*(1), 55-59.
- Owen, D. W. (2007). In Search of Excellence in Profession of counseling; (R. Özyürek, F. Korkut-Owen and D. W. Owen, Ed). Pathways to Professionalism in Counseling (pp. 1-20). Ankara: Nobel Publishers.
- Özaydın, A. (2001). *Resmi ilköğretim okullarında yönetici, rehber öğretmen ve sınıf öğretmenlerinin görüşlerine göre rehberlik uygulamaları* [Guidance services in elementary schools according to principals, counselors and teachers]. Unpublished master's thesis, Yıldız Teknik University, İstanbul, Turkey.
- Petersen, R. (2008). *The structure of the educational system in Costa Rica*. Retrieved November 13, 2008, from www.costaricalaw.com/index.php?option=com_content&task=view&id=45&Itemid=38
- Peterson's Graduate Planer: Student Edition (2007). Retrieved February 26, 2007, from www.petersons.com/graduate_home.asp?path=gr.home
- Ponec, D. L., & Brock, B. L. (2000). Relationships among elementary school counselors and principals: A unique bond. *Professional School Counseling, 3*(3), 208-218.
- Söker, V. (2007). *İlköğretim öğretmenlerinin sınıf yönetimi konusunda rehberlik servisinden beledikleri: Zonguldak Örneği* [Expectations of elementary school's teachers for class management from guidance services (The sample of Zonguldak)]. Unpublished master's thesis, Gazi University, Ankara, Turkey.
- Taylor, K. R. (2002). Do you know what your guidance counselor should know? *Principal Leadership, 2*, 59-62.
- Turkish Statistical Institute. (2007). *Adrese dayalı nüfus kayıt sistemi 2007 nüfus sayımı sonuçları* [Results of 2007 consensus]. Başbakanlık, Türk İstatistik Kurumu Haber Bülteni. Sayı: 9, 21 Ocak 2008.
- Turkish Statistical Institute. (2008). *Eğitim kurumlarının kademelere göre okul, öğrenci ve öğretmen sayısı* [School, students and teacher numbers related to educational level 2007-2008]. Retrieved December 11, 2008, from www.tuik.gov.tr/VeriBilgi.do?tb_id=14&ust_id=5.
- United States Census Bureau. (2008). *2008 U.S. POPClock projection*. Retrieved December 7, 2008, from www.census.gov/population/www/popclockus.html
- United States Department of Education. (2008). *Structure of U.S. education*. Retrieved October 27, 2008, from <http://www.ed.gov/about/offices/list/ous/international/usnei/us/edlite-structure-us.html>

- United States Department of Education, National Center for Education Statistics. (2008). *Digest of education statistics: 2007*. Retrieved December 13, 2008, from www.nces.ed.gov/fastfacts
- United States University Directory. (2008). *Education encyclopedia: Global education reference Costa Rica history and background, constitutional and legal foundations, educational system: Overview, preprimary and primary education, secondary education*. Retrieved October 2, 2008, from www.education.stateuniversity.com/pages/305/Costa-Rica.html
- United Nations Development Programme. (2008). *2007/2008 Human Development Report: Turkey*. Retrieved December 13, 2008, from www.hdrstats.undp.org/countries/data_sheets/cty_ds_TUR.html
- Zalaquett, C. P. (2005). Principals' perceptions of elementary school counselors' role and functions. *Professional School Counseling*, 8(5), 451-457.

Okul Psikolojik Danışmanları ile Yöneticilerin Üç Ülke Örneğinde İşbirliğine Dayalı Ortaklıkları

(Özet)

Yüz yıldan daha fazla bir zamandır okulların yapılarına ve örgütlenmelerine bakış açısında ciddi değişiklikler olmaktadır. Bu değişikliklerden birisi ve belli ki de en yaygın olanı okullardaki okul psikolojik danışmanlığının eğitimle bütünleşmesidir. Orijinalde "rehberlik" olarak tanımlanan kavram ilk önce Amerika Birleşik Devletleri'nde (ABD) ortaya çıkmış olsa da gelişmiş ülkelerde geniş olarak kabul edilmiştir. Okulların sadece bilgileri öğretme yerleri olmaması, aynı zamanda öğrencileri gelecekte üretken, yetişkin dünyasına daha anlamlı katkıları olan kişiler olarak hazırlayan yerler olmaları gerektiği fikri yaygınlaşmıştır. Bu fikir, okullarda, okul psikolojik danışmanlığı ve rehberliği programlarının eğitimin bütünleyici bir parçası olmalarına yol gösterici olmuştur. Okul psikolojik danışmanları eğitim sistemine, bir çok ülkelerde, sınıf ve alan öğretmenleri ve eğitim yöneticilerinden sonra üçüncü ana öğe olarak katılmıştır.

Bu makalede, Kosta Rika, Türkiye ve ABD'de okul psikolojik danışmanlığının gelişimi, statüsü, uygulanması, yönetsel açıdan ele alınışı, okul yöneticileriyle okul psikolojik danışmanlarının ilişkisi gözden geçirilerek uluslararası bir bakış açısıyla sunulmaktadır. Büyük tarihsel, kültürel ve politik farklılıklarına rağmen, karşılaştırılan bu üç ülkede okul psikolojik danışmanlığı kendi eğitim sistemlerine kendilerine özgü yollarla bütünleşmiştir. Türkiye ve Kosta Rika'da, tüm eğitim sistemi yüksek derecede merkezileşmiş bir yapı gösterirken ABD'de eyaletlerin ve bölgenin gereksinimlerine ve koşullarına göre farklılaşan yerelleşmiş bir eğitim sistemi sergilenmektedir.

ABD'de psikolojik danışmanlık ve rehberlik, 20. yüzyılın başında tamamen ülkenin sosyal, eğitimsel ve benzeri gereksinimleri dikkate alınarak ve mesleki rehberlik olarak başlayan bir harekettir. Kosta Rika'da 1956'da

Kosta Rika Üniversitesi'nin ABD'nin etkisiyle Psikolojik Danışma Lisans Programı açması ile başlayan psikolojik danışmanlık ve rehberlik, Türkiye'de de Amerika Birleşik Devletleri'nden gelen uzmanlarının önerilerinin Milli Eğitim Bakanlığı tarafından kabul edilmesiyle 1950'lerde başlamıştır. Dolayısıyla psikolojik danışma ve rehberliğin ilk ülkede yüz yıldan daha uzun bir geçmişi varken diğer ülkelerdeki geçmişi yarım yüzyıldan biraz fazladır.

Gelişimsel olarak ele alındığında da psikolojik danışma alanının daha önce başlamış olmasının da etkisiyle mesleki olarak tanımlanmasının, kabul edilmesinin, psikolojik danışman eğitiminin standartlaştırılmasının ve kalite güvencesi çalışmalarının en fazla ABD'de geliştiği görülmektedir. ABD'de sınıf öğretmeni iken, en az master derecesi olarak okul psikolojik danışmanı olunabiliyorken, diğer iki ülkede psikolojik danışman olabilmek için yaygın olarak lisans düzeyinde eğitim verilmektedir. Türkiye'de lisans eğitiminin yanı sıra yüksek lisans ve doktora düzeyinde de psikolojik danışma eğitimi verilirken, Kosta Rika'da lisanın yanı sıra sadece lisans üstü düzeyinde eğitim verilmektedir.

Okul psikolojik danışmanlığının yönetsel düzeyde tanımlanması da üç ülkede farklılık göstermektedir. Şöyle ki, ABD'de daha belirgin olan psikolojik danışmanların rol ve işlevlerinin, Türkiye'de biraz daha belirsiz, Kosta Rika'da ise en belirsiz olarak tanımlandığı görülmektedir. Her üç ülkede de psikolojik danışmanlığın en önemli görevlerinin öğrencilerin akademik, sosyal/kişisel ve mesleki alanlarda gelişimlerini desteklemek olduğu üzerinde durulmaktadır. Aslında bu konular, öncelikle Amerikan Psikolojik Danışma ve Rehberlik Derneği'nin üzerinde durduğu gelişimsel alanlar olması nedeniyle önce ABD'de vurgulanmış, ardından diğer ülkelerde de benimsenmiş gibi görülmektedir.

Okullarda görece yeni ve kendine özgü rol oynasa da okul psikolojik danışmanlığı en fazla okul yöneticilerinden etkilenmektedirler. Bunun en önemli kaynaklarından birisi okul yöneticilerinin her üç ülkede de okuldaki her tür etkinliğin yanı sıra psikolojik danışma ve rehberlik servislerinin de ana sorumluları olmalarıdır. Okullarda, yıllar içinde çok iyi tanımlanmış rol ve sorumlulukları olan, öğretmen ve yöneticilerin, rolleri hala tanımlanma süreci içinde olan okul psikolojik danışmanlarla olan ilişkilerinin niteliği okulun verimliliğini etkileyen önemli unsurlardandır. Okul psikolojik danışmanlarının rol ve işlevlerinin, özellikle okul yöneticilerince nasıl algılandığı, psikolojik danışma ve rehberliğin işlevlerini okul için ne kadar önemli buldukları konusu bu anlamda büyük önem taşımaktadır. Yapılan çalışmalar Türkiye'deki okullardaki yöneticilerin psikolojik danışma ve rehberliğin rol ve işlevlerini algılamaları ile ilgili farklı sonuçlar göstermektedir. Şöyle ki okul yöneticilerinin psikolojik danışmanların rol ve görevlerini iyi bilmediklerini ve onlara görevleri dışında işler yüklediklerini gösteren araştırma sonuçlarının yanı sıra yöneticilerin psikolojik danışmanlığı çok destekleyen tavırlar içinde olduğunu gösteren araştırma sonuçlarına da rastlanmaktadır. Okul yöneticilerinin psikolojik danışma ve rehberlik hizmetlerini, okulun misyonunun gerçekleştirilebilmesi için çok önemli

bulduklarını ve bu hizmetin gerekliliğini vurguladıklarını gösteren araştırma sonuçları da vardır. Genel olarak bakıldığında okul yöneticilerinin bu hizmeti daha önemli bulduklarına ilişkin araştırma sonuçlarına daha fazla rastlanmaktadır. Türkiye’de psikolojik danışmanların okul sistemindeki yerleri yöneticiliğe ya da öğretmenliğe yakın olarak değil, daha belirsiz bir statüde algılanmaktadır. Kosta Rika’da ise okul yöneticileri, psikolojik danışmanları bir bakıma müdür yardımcısı gibi algılamakta ve yönetsel işlerin yapılmasında onlardan destek almaktadır. O nedenle psikolojik danışmanlar, yönetime daha yakın bir statüde algılanmaktadır. Bu tip bir yapılanma nedeniyle yöneticilerle psikolojik danışmanlar arasında mesleki roller konusunda fazla çatışmaya rastlanmamaktadır. Amerika Birleşik Devletlerindeki araştırmalarda da psikolojik danışmanların yönetsel işlere daha fazla katkıda buldukları biçiminde algılamaları olduğunu gösteren çalışmaların yanı sıra son yıllarda bu iki grubun psikolojik danışmanların rol ve işlevlerini benzer biçimde algıladığını gösteren çalışmaların sayısı daha fazla gibi görünmektedir. Bu iki meslek grubunun etkili bir işbirliği içinde çalışmaları halinde okulun başarısının arttığı ve okulun işlerini rahatlıkla yerine getirildiği üzerinde duran araştırma sonuçları da vardır. Okul sistemi açısından bakıldığında ise psikolojik danışmanlar öğretmenlerden biraz daha fazla, yöneticilerden daha az statülü olarak ele alınmaktadırlar. Sonuç olarak her üç ülkede de psikolojik danışma ve rehberlik hizmetlerinin etkili olarak verilmesinde okul yöneticilerinin rolü çok büyüktür. Yapılan çalışmalar psikolojik danışmanların ve okul yöneticilerinin birbirlerinin rol ve işlevlerini iyi anladıkları ve işbirliği içinde oldukları bir okul ortamının daha etkili olabileceği göstermektedir. Dolayısıyla psikolojik danışman ve eğitim yöneticisi yetiştiren programlarda, her iki gruba da diğer grubun rol ve işlevlerinin daha ayrıntılı verilmesi ve nasıl işbirliği yapılabileceğinin üzerinde durulması anlamlı görülmektedir.

Anahtar sözcükler: Okul psikolojik danışmanlığı, eğitim yöneticileri, Kosta Rika, Türkiye ve ABD

A Study of Graffiti in Teacher Education

S.Nihat Şad*
Mustafa Kutlu**

Suggested Citation:

Şad, S.N., & Kutlu, M. (2009). A Study of graffiti in teacher education. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 39-56.

Abstract

Problem statement: Graffiti is about self-expression. When youth cannot find people to listen to them, they may express their strongly felt, internal experiences and emotions safely by writing on public property. Thus, graffiti can be handled as a counseling issue. When this self-expression of a thought, wish, or attitude comes from prospective teachers, the difficult work of sorting these issues out may help us develop better teacher-education programs and produce better teachers. Thus, this work takes the issue of graffiti by prospective teachers as an interdisciplinary issue, bridging counseling and teacher training.

Purpose of the Study: This research aims at extending the understanding of the contents of and underlying reasons for graffiti written by prospective teachers.

Method: This study is based on the content analysis method and supplemented with structured interviews. A total of 178 inscriptions were analyzed. In addition, six graduating students were interviewed.

Findings and Results: One of the major implications of this qualitative research was the presence of some methodological problems inherent in the graffiti-related studies. Second, it was found that bathroom graffiti might point to the importance for student teachers to express themselves in terms of sexual, political, and religious issues and to socialize through proper communication with others. Finally, the results of the analysis indicated that there were two main spheres of graffiti production: labs/classes and restrooms. The first group included rather socially acceptable topics, which focused on the need for belongingness, homesickness, romance, and humor or the form of someone's name and signs (doodling). The second group included more anonymous inscriptions, mainly about sex and politics/religion. In this category, men were found to write more than women did.

Conclusions and Recommendations: The findings suggested that the anonymous nature of bathroom graffiti particularly makes it difficult to be inquired about through interviews. Prospective teachers' need to express themselves and socialize can be further met through curricular and/or extracurricular activities

* PhD student, İnönü University, Faculty of Education, Curriculum and Instruction, snsad@inonu.edu.tr

** Assoc. Prof. Dr., İnönü University, Faculty of Education, Counseling and Guidance, mkutlu@inonu.edu.tr

providing opportunities to communicate and interact socially. In addition, graffiti in labs/classes points to the possibility that the faculty's failure to involve the students actively in class causes boredom for students, which implies a need to apply more participative instruction on the part of instructors. As for the bathroom graffiti, the finding that men find more to say about sex, politics/religion, and humor, while women do not even write about romance may indicate a deprivation of liberal perspectives on the part of women.

Keywords: Graffiti, prospective teachers, counseling, and teacher education

How do we interpret the graffiti written by prospective teachers? Do we really need to know the context and reasons for what they write about on the desks and walls of their classrooms? Some may not see graffiti as a serious field of study (Gadsby, 1995). However, as teacher trainers and parents of those children who are to be educated by prospective teachers, we are curious and wish to discover the answers to these questions. Each graffito expresses a thought, wish, or attitude (Gach, 1973). As we go through the difficult work of sorting them out, we learn more about the thoughts, wishes, and attitudes of future teachers, which can help us to develop better teacher-education programs and produce better teachers.

When youth cannot find people to listen to them, they may express themselves by writing on public property (Kurt, 2002). Even when they find people with whom to talk, they still might hesitate to express themselves with these individuals and choose graffiti as a form of reflecting strongly felt, internal experiences and emotions safely (Klingman, Shalev, & Peariman, 2000). Thus, graffiti is also about self-expression (i.e., expression of emotions) and as such is a counseling concern. Graffiti then takes on a more critical direction given that the participants in this study are future teachers who will be engaged professionally with such questions like self-expression in their work with students. Thus, graffiti by prospective teachers is an interdisciplinary issue, bridging counseling and teacher training. This research extends the understanding of the contents of and underlying reasons for graffiti written by prospective teachers.

Defining Graffiti

Gach (1973) defines the Italian word *graffiti* (plural of *graffito*) as "statements and drawings...penciled, painted, crayoned, lipsticked, or scratched on desk and walls, particularly restroom walls" (p. 285). The Office of Juvenile Justice and Delinquency Prevention in the U.S. Department of Justice (1998) describes graffiti in criminal terms as vandalism, which is "a willful destruction or defacement of private or public property" (p. 2). Bates and Martin (1980) think that the anonymous messages of graffiti are socially uncontrolled manifestations of thought. Stocker, Dutcher, Hargrove, and Cook (1972) look at graffiti from a cultural perspective and according to them graffiti shows "the patterns of customs and attitudes of a society" (p. 356). Kan (2001) sees graffiti in bathrooms/lavatories as "a silent and mindless protest against the large educational system that alienates students' primary needs" (p. 19). Although graffiti is considered a serious and costly act of vandalism by many societies (Klingman, et al., 2000; Walsh, 1996), it is also acknowledged as an art form by some (Kan, 2001; Richardson, 1999). This is mainly because certain graffiti features "novelty in both language and visual representation with a unique and holistic aesthetics" (Kan, 2001, p. 21).

Who Writes Graffiti: Reasons

Literature on the motivations for graffiti is quite rich. Although this act of writing or drawing on public walls as a form of anonymous self-expression is a common event (Bates & Martin, 1980; Kurt, 2002), the performers of graffiti are usually suppressed individuals in society (Kan, 2001; Richardson, 1999), who express themselves by making their mark on society or who enjoy the pleasure of risk taking (Richardson, 1999). Blume (in Gadsby, 1995) identifies some of the general motives for writing graffiti, including furnishing proof of one's existence; acting on a need to express oneself; documenting one's membership in a group; taking pleasure in aesthetic, creative, and physical acts; boredom; expressing criticism, protest, rejection, and agreement; marking out territories; and searching for contact with others.

As for the educational context, Gottfredson and Gottfredson reported that most of the graffiti in schools are produced by students "who do not like school, whose school performance is poor, and who are not committed to education" (as cited in Center for Mental Health in Schools at UCLA, 2007 in the preface). An early study (Gach, 1973) argued that students under pressure tend to create graffiti, thus communicating attitudes and feelings that they would hesitate to utter publicly. More specifically, middle-class teenagers or pre-teen boys—most of whom are not successful at school or in sports—reportedly create most of the street-type wall graffiti (Richardson, 1999). Kan (2001) interprets graffiti as an unconscious rejection of the kind of learning that is not helping students construct personal meanings and effectively integrating their inner needs to promote growth. As a matter of fact, Flaherty reports (as cited in Kan, 2001) that there is less or no such vandalism in "schools that manage a successful community of learning" (p. 19).

What is Written: Content

Graffiti content helps us see the character of a society (Gadsby, 1995). It is an accurate gauge of the socialization process (Bates & Martin, 1980). Some investigators study the content of graffiti as a nonreactive indicator of human attitudes, thoughts, wishes, or attitudes (Bates & Martin, 1980; Gach, 1973). With a more gender-based approach, while the walls in men's restrooms are reported to have usually more sexual content than those in women's restrooms, especially on homosexuality, the walls in women's restrooms have more romantic writing (Green, 2003; Schreer & Strichartz, 1997; Sechrest & Olson, 1971). According to Bruner and Kelso (1980), women's graffiti is more communicative; that is, when a question is posed via graffiti, it is answered in a serious manner. Green (2003) reports women's restrooms include more arguments on religion and philosophy, whereas men's most common themes are sex and politics.

Purpose of Study

The purpose of this study was to investigate the graffiti written by prospective teachers in terms of their content and the possible reasons for writing them.

For this purpose, answers to the following research questions have been sought:

1. What are the thematic categories of the graffiti written by prospective teachers?
2. What can be the possible reasons for the prospective teachers to write graffiti?

Method

Research Design

This study was based on qualitative design. The method of content analysis was used to glean a rich understanding about the graffiti written by the prospective teachers, which was supplemented with the analysis of the data from structured interviews with six prospective teachers. Qualitative content analysis aims at providing knowledge and understanding of the phenomenon under study (Downe-Wamboldt, 1992). Content analysis can be used in qualitative research when direct observation or interviewing is impossible (Yıldırım & Şimşek, 2003). In this study, the impossibility of direct observation was the primary reason for choosing content analysis because writing graffiti is an anonymous practice; that is, the producer of the graffiti is generally unknown. Yet, Biklen and Casella (2007) suggest that textual analysis should also include interviews to analyze the perspectives of the target audience, namely the prospective teachers in this study. Thus, content analysis was coupled with structured interviews with six students for the sake of triangulation.

Sample

The university where the data were collected has two buildings (Blocks A and B) where students in the college of education are hosted. Because Block B was recently taken over from the department of engineering, it was excluded from the study. A total of 178 inscriptions, the sample for this study, were collected from five men's restrooms, five women's restrooms (each with four stalls), four laboratories, and thirty-three classrooms with a capacity of about 1,000 students. Block A has been hosting the students from the college of education since 1997, but most inscriptions collected were assumed to be recent since the building was cleaned and repainted in 2006. It is also assumed that during these two years, many of the inscriptions could have been cleaned by the staff or deleted by other students. Inscriptions were either written in pen or pencil or scratched. For this study, the entire building, including the restrooms and classes/labs, were scanned for graffiti content, and only 178 legible ones were found and considered for the study. Although most of the graffiti sample was in the form of a word or sentence, 21 included signs (doodling).

It is assumed that prospective teachers produced the analyzed graffiti, since the building is primarily allocated to the students of the college of education. In addition, the academic and administrative staff has their own offices and restrooms in the adjacent building. Thus, it did not seem likely that faculty or administrative staff would use students' restrooms or that other students from other faculties, who are housed in different buildings, would use the same labs/classes and restrooms.

As for the interviews, participants included six graduating male students from different departments within the college of education who admitted having written graffiti at least once during their education. This sample was one of convenience, obtained from among 36 graduating students (11 women, 25 men) who visited the researcher's department office during the summer either for the purpose of returning the robes and caps they wore at the graduation ceremony or for seeking signatures

for graduation purposes. Only 6 of the 36 students admitted having written graffiti and were, therefore, deemed appropriate to include in the study.

Procedure

The graffiti was recorded in two repeated stages: First, two senior students, one male and one female, collected data using digital cameras during the 2007-2008 spring term for a seminar course. Second, the researchers re-examined the entire building at the conclusion of the term after the building was vacated for summer vacation. The researchers wrote down the graffiti as a check against the graffiti collected by the student assistants. As a result, 97% agreement was achieved between the graffiti recorded in the two stages.

The interviews with the six students were brief. After acknowledging they had written graffiti sometime during their teacher training, they were asked three simple questions: where, what, and why did they write? Their answers were noted by the first researcher. It can be assumed that their answers were reliable since they were about to leave the institution as graduates. On the other hand, as a limitation of the study, the anonymous nature of graffiti (Bates & Martin, 1980) suggests that both the initial 36 and final 6 interviewees may have preferred to withhold or disguise some of the data regarding the place and content of their possible graffiti experience due to personal embarrassment or perceived standing with the faculty member.

Data Analyses

For the analysis of the inscriptions, Nvivo 8 software program was used. Nvivo is a practical qualitative software program, which is mainly used for data analysis in grounded-theory type of studies (Kuş, 2006, p. 39). Yet, in this study, it was used for the sake of more systematic coding and categorization. Both researchers independently coded all of the data with 89% agreement between coders. Most of the disagreement was due to graffiti that was either not categorized (placed in a *miscellaneous* category) or graffiti that fell into multiple categories. The final distribution of the categories with thematic details is displayed in Table 1. A similar procedure was followed with the data obtained in the interviews, which revealed 100% agreement in the coding of categories and themes (see Table 2).

In the findings and discussion sections of this study, some of the graffiti and statements from the interviews are directly quoted. For this purpose, their English translations are provided. In order to validate the translations of the graffiti and interview statements, two independent bilingual academicians, one from the counseling and guidance department and another from the English language and literature department, were asked to review the translations made by the researchers. Based on their remarks, translations were revised to achieve an adequate degree of equivalencies in translation (Beaton, Bombardier, Guillemin, & Ferraz, 2000).

Table 1
Thematic Distribution of Graffiti.

CATEGORY	PLACE			Total
	Lab/Class (public)	Restrooms (non-public)		
		Men	Women	
Total number of graffiti	128	34	16	178
POLITICAL/RELIGIOUS				
Political Religious				
In favor	1	3	1	4
Against		1	1	2
Pure religious	2	1		3
Apolitical		1		2
Terrorist propaganda		1		1
Bourgeois criticism		1		1
Subtotal	3	8	2	13
COMMUNICATION				
Search for contact	1		4	5
Dialogue				
Receiver known	8			8
Receiver unknown		3	2	5
Whoever reads	3	7		10
Subtotal	12	10	6	28
HUMOR	9	5	2	16
Subtotal	9	5	2	16
SEXUAL				
Homosexual	1	5	3	9
Anti-homosexual		4	1	5
Assault		2	2	4
Eulogy	1			1
Subtotal	2	11	6	19
NEED FOR BELONGING/ HOMESICK				
City/hometown				
Family	32	2		34
Team			1	1
Desire to leave	4	1		5
	4			4
Subtotal	40	3	1	44
ROMANCE				
Declaration of love	8			8
Sorrow of love	4			4
Definition of love		2	1	3
Platonic love	2			2
Subtotal	14	2	1	17
PHILOSOPHICAL				
Subtotal	3	3		6
OTHERS				
Names	23			23
Signs(doodling)	21			21
Miscellaneous	9		2	11
Subtotal	53		2	55

Note. Adding the subtotals does not give the total number of graffiti, since some graffiti fall into multiple categories.

Table 2
Graphical Display of Interview Results

Interviewee	Where?	What?	Why?
A	Class	Name	Boredom
B	Class	Sign (doodling)	Boredom
C	Class	Sign (doodling)	Boredom
D	Class	Sign (doodling)	Boredom
E	Class	Romance	Boredom, love
F	Class	Communication	Boredom

Findings and Results

Numbers and Topics

Of the 178 inscriptions, a majority (71.9%) were in the labs or classes – mainly on desks and a few on walls – whereas the remaining 50 (28.1%) were in the restrooms, mainly on walls and doors. It is difficult to decide whether this amount of inscription is low, high, or average given the time period (estimated 2 years), number of students ($N=4,488$), and the assumption that many of the inscriptions might have been cleaned. However, the ratio of students (4,488) to the number of inscriptions (178) over 2 years reveals that 1 out of 12 students writes an inscription each year, assuming that each person writes only one inscription. The overrepresentation of graffiti in classes and labs compared to the restrooms can be attributed to the time spent in both locations. Based on the interviews with the students (see Table 2), it seems that all the inscriptions in classes/labs were written during a lesson because of boredom and lack of interest in the lesson:

1. When you become bored during the lesson, you unwittingly inscribe something on the desk. [Interviewee C]
2. When I became bored as the teacher or one of the classmates was giving a presentation, I wrote a message on the desk since there was no paper... [Interviewee F]

A thematic analysis of the overall graffiti in the classes and labs shows that the contents are more diverse than the ones mentioned in the interviews. The most evident finding is city/hometown names ($n=32$; 25%), personal names ($n=23$; 17.9%), and signs (doodling) ($n=21$; 16.4%), followed by romance ($n=14$; 10.9%), and other minor categories.

Gender

As for gender, compared to the general population of students (2,080 men + 2,408 women = 4,488), men tended to write more ($n=34$; 68%), whereas women wrote less ($n=16$, 32%) in the restrooms. This finding was in consistency with the research by Stocker et al. (1972), who found that female students on college campuses wrote less

graffiti. What is remarkable is that female underrepresentation can be associated with the conservative nature of the university (Green, 2003). As a matter of fact, a recent study (Aksu et al., 2007) on 18,226 prospective teachers from 51 universities in Turkey revealed a generally conservative profile. These conservative attitudes were also evident in some of the graffiti:

3. God help all of us! [Class 402]

4. Islam. [Class 201]

5. Man lives his fate/Whatever it brings...Only the *** is invulnerable... [Men's Restroom] (the word God/Allah was deleted most likely as its mention in a restroom is considered a sin)

Examining the distribution of the topics by gender, it seems men's restrooms are dominated by sexual content ($n=11$; 32.3%), followed by a form of communication ($n=10$; 29.4%), political/religious content ($n=8$; 23.5%), and humor ($n=5$; 14.7%). As for prospective female teachers, they also wrote the most, though not as much as their male counterparts, about sex ($n=6$; 37.5%), usually in the form of communication ($n=6$; 37.5%). These findings seem to agree with relevant literature, which suggests that both men and women write more about sex, dominated by homosexuality, that men also write more about politics (Green, 2003), and that women's graffiti are more interactive (Bruner & Kelso, 1980). It is also remarkable that romance is one of the least identified contents in women's restrooms contrary to the relevant literature (Bates & Martin, 1980; Green, 2003; Workman et al., 1983).

Thematic Categories

Political/Religious. A total of 13 graffiti with political/religious content were represented in the men's restrooms ($n=8$; 61.5%) with themes including favoring or opposing a political/religious viewpoint [The lamb (*symbol of the ruling party*) is fading out; One day the universities will *** (illegible)... this infidel way of life will end...], pure religious wish (see 5), a political message [It is neither left nor right, football is the rightest choice (*meaning the most reasonable*)], terrorist propaganda, and bourgeoisie criticism [...A happy minority which muzzled our past attempts to govern our country...]. In fact, the religious/political content of the sample is not concentrated on a single sect, and opposing political/religious views are communicated through an ersatz bulletin board:

6. If showing skin means being civilized, then the animals are more civilized than you. N.F.Kısakürek [Men's Restroom].

7. Get a life! Do make progress! [In response to 6].

8. Does one need to show a part of her body to get progress [In response to 7].

The issue of women covering their heads or the way they dress has always been a delicate issue in the Turkish secular republican regime. What seems paradoxical is that this religious/political issue was discussed by men in the men's restroom. This can be a result of Turkey's patriarchal culture in which women's behavior is something about which men have a say.

Communication. Though not exactly a topic, but rather a form or purpose of writing graffiti, communication stands as another important category. Although the

content of the communication may range from humor (see 19) to searching for a sex partner (see 31), the analysis focused on the possible form and purpose of communication. Communication, especially in the form of dialogue whose receiver is possibly known ($n=8$; 28.5%), was evident in classes:

9. Fortunately! [Class 405] (*A possible response in various contexts in Turkish*)
10. How sly you are! [Class 402]
11. He wouldn't get angry! [Class 402] (*A possible answer or advice*)
12. Let's write to each other! Take out a piece of paper! [Class 203]
13. Me, too! [Class 203]
14. Hey dude, why the hell did we come to this department. [Class 203]
15. I am hungry!.....Me, too! [Class 303]

These findings are supported by one of the interviewees (Interviewee F), who revealed that one reason prospective teachers write on their desks is to communicate with a classmate during class. A possible reason for these dialogues, given the interviews and the contents of the inscriptions, can be long lesson hours without breaks, unattractive and tiring methods of instruction, such as narration, or failure to engage the students.

These dialogues happen to change, especially in men's restrooms, into a response to someone the student does not know (see 6, 7, and 8), an anonymous statement (see 16), a derogatory comment (see 17), or a curse aimed at whomever reads the inscription (see 18):

16. I hope we all graduate without any trouble. It is bad to end something but we have a new life ahead. I wish success in your lives! [Men's Restroom]
17. Are you both shitting and smoking at the same time? [Men's Restroom]
18. I f*** all of you! [Men's Restroom]

Humor. The category of humor includes graffiti in which participants attempted to express their sense of humor to the reader, which is quite easy to distinguish from the other inscriptions. Moreover, humorous graffiti in the restroom and class/lab are quite distinctive. Whereas restroom graffiti are more likely to mock the reader (17, 19, and 20) and refer to a popular movie (21) or to the absurd (22), the majority of the remaining class/lab graffiti (9/16) are organized in the form of a wordplay or logical proposition (23, 24, 25, 26, and 27):

19. (*on the right wall*) Look at the left wall!... (*on the left wall*) Why are you looking? Fool! Keep on shitting!!! [Men's Restroom]
20. (*on the right wall*) Look at the left wall!...(*on the left wall*) Look behind!....(*On the rear wall*) You fool, you fool!!! [Women's Restroom]
21. Free for the Landowner! [Women's Restroom]
22. Wow! Could I ever imagine that I would shit here one day? You can't know how very impressed I am now! [Men's Restroom]
23. Once there was "one" of a man (both means *someone* and *a man with a low mark in exam*), then he fixed it in the second term. [Lab]
24. If working were something good, they wouldn't pay for it. [Lab]
25. As you happen to be born to this world, study physics and chemistry like mad :) [Lab]

26. I used to be arrogant, but now I am perfect. [Class 201]

27. I bet I will quit gambling. [Class 203]

Sexual Content. Sexual content seems to be dominant in men's (11/19) and women's (6/19) restrooms, whereas labs/classes contain less graffiti on sexual discourse (2/19). Among the specific contents, homosexual messages were found to be overrepresented, which is consistent with the literature (Green, 2003). Accordingly, five of the scripts in men's restrooms (e.g., 28, 29, and 30) and three in women's restrooms (e.g., 31) were in favor of homosexuality. A remarkable finding was the presence of five other messages (32, 33, 34, and 35) strongly opposing the homosexual scripts noted above. Apart from the homosexual or anti-homosexual graffiti content, the remaining four sex-related scripts in restrooms (two in men's and two in women's restrooms) were found to be more heterosexual oriented, taking the form of sexual assault (18 and 36). The other two inscriptions insulting/assaulting to heterosexuals in women's restrooms were partly wiped clean (especially the most indecent words).

28. Freedom for gays. [Men's Restroom]

29. We demand a Gays and Lesbians Community at Inonu University. [Men's Restroom]

30. I need a big c***! [Men's Restroom]

31. Your fantasies are important to me. Call 0505*****. [Women's Restroom]

32. Death for gays! [In response to 28]

33. You can get over it. Make sure that you believe you are normal! [In response to 29]

34. A man wouldn't be f***ed in the ass! [In response to 30].

35. I hope they f**k you! [In response to 31].

36. Do you have your mother? [Men's Restroom]

Based on the findings in this category, it seems that although homosexuality receives a relatively high level of coverage among sexual graffiti, it actually receives sharp opposition and is not welcome. In interpreting this finding, we refer to Kinsey et al. (as cited in Sechrest & Olson 1971), who found a relation between low socioeconomic status and graffiti with anti-homosexual and rather heterosexual content. Since it was reported in a recent study (Aksu et al., 2007) that the prospective teachers studying at colleges of education in Turkey come from families with low to moderate socioeconomic status, similarly, in our study, the low socioeconomic status of the participants may be the reason for the anti-homosexual and strong heterosexual graffiti content. The findings of Stocker et al. (1972) further suggest that this intolerant environment can give rise to the production of graffiti with homosexual content because they found an absence or at least a decrease in such graffiti in societies with more tolerant attitudes towards homosexuality.

Need for Belonging/Homesickness. The inscriptions under this category were perhaps the easiest to code, yet the most difficult to comment on. The difficulty stemmed mainly from the city/hometown names ($n=34$), which constituted the most

popular specific content, especially among lab/class graffiti, with 32 entries on desks only. Although one reason for writing them on desks can be general boredom during a lesson as suggested by the interview data, the preponderance of city/hometown names deserves discussion. People have some attachment either personally or collectively to their locality ranging from village, town, city, or nation (Knight, 1982). This can be traced to the feeling of belonging to a group who will provide security against external threats. The presence of names of some Turkish local (e.g., Batmanspor and Trabzonspor) or national football teams (e.g., Galatasaray and Fenerbahçe) ($n=4$) also seems to be in agreement with the same need to identify with a group.

Still another possible reason for inscriptions of city/hometown names could be students' feelings of homesickness (see 37). This interpretation is supported by the presence of two related topics: family ($n=1$) (see 38) and the desire to leave ($n=4$) (see 39 and 40).

37. Amed...my hometown, I have been dreaming of you! [Men's Restroom]

38. This is my family (*an accompanying picture illustrating parents hand in hand and two children*) [Women's Restroom]

39. 22 for dawn (*number of days before military service is over*) [Class 405]

40. I hate the school [Class 203]

Romance. Although romance was one of the least referred contents in women's restrooms, contrary to the literature (Bates & Martin, 1980; Green, 2003; Schreer & Strichartz, 1997; Sechrest & Olson, 1971), its presence in classes/labs cannot be underestimated ($n=14$). Prospective teachers may prefer to write about love publicly, unlike political/religious and sexual content. They especially prefer to declare their love on desks ($n=8$), either by writing (see 41 and 42) or drawing/carving hearts containing initials. In addition, in classes, they express their feelings of grief ($n=4$, see 43, 44, and 45) or platonic love ($n=2$, see 46). One interviewee stated that he once wrote down the name of his lover in fine shorthand (see 47):

41. I damn love! [Class 202]

42. I love! [Class 405]

43. S/he had not loved me! [Class 405]

44. I heard the girl I love got married, she belongs to somebody else now [Class 405]

45. What if I said "Come back to me?" [Class 203]

46. I love you from a distance, you do not even know. [Lab]

47. I was already not interested in the lecture...I found myself writing my girlfriend's name on the desk. [Interviewee E]

Contrary to the categories above, prospective teachers preferred to define love in restrooms:

48. Love is not the *** (*illegible*) on the lips, but tears in eyes. It is to find the lover as *** (*illegible*) dying. [Men's Restroom]

Philosophical Content. Although the literature (Bates & Martin, 1980; Green, 2003) reports philosophy is a popular content among women, in the present study, no philosophical expression was found in women's restrooms, and only three such graffiti were found in classes/labs. All 6 items (3 in men's restrooms and 3 in classes/labs) were philosophical in nature, mostly favoring the message *carpe diem*, whereas some took a humorous turn (49 and 50), and others were more straightforward (51):

49. Life is between 3.5 and 4, so you either turn 3.5 (a *colloquialism meaning to get afraid*) or you live 4/4 (a *colloquialism meaning perfect*).[Lab]

50. If working had been something good, they would not pay you for it.[Lab]

51. Don't think about the past, since it's passed, you cannot change it...And you cannot know about future, thus live for the moment.[Men's Restroom]

Others. Some graffiti were unable to be categorized. They constitute, however, a considerable share in the total (55/178). Although 23 of these were names, it is difficult to understand whose names they are. Given that they exclusively appear in classes/labs ($n=23$) and given the statement of one of the interviewees (Interviewee A), the reason for writing these names can be boredom during lessons, especially considering the signs (doodling including initials, stars, cubes, triangles, flowers, zigzags, and parallel lines) ($n=21$). This reason was endorsed by Interviewees B, C, and D who stated they had once doodled on the desks during a boring lesson. The other 11 were categorized as miscellaneous since they were independent of most of the categories.

Conclusions and Recommendations

In this section, the findings of the study are discussed, and some conclusions are drawn. The major implications of this qualitative research center on methodological problems are inherent in the graffiti-related studies: (1) graffiti as vandalism, (2) implications for teacher training, and (3) what lies behind public and private graffiti written by prospective teachers from a counseling perspective.

Like other graffiti researches (Green, 2003), this study suffers from some methodological problems. First, the sample graffiti were assumed to be written by prospective teachers as the data were collected from the building of the college of education. Yet, other students from other faculties might have written some of the graffiti, and also, some men might have written in women's restrooms or vice versa. Second, individuals that write bathroom graffiti usually hide their identities, which Bates and Martin (1980) calls anonymity, because such activity involves vandalism and some degree of indecency; this fact might have hindered researchers' effort to collect correct qualitative data. In particular, the data collected through interviews may not be complete, credible, or reliable because students were able to maintain their anonymity. That is very evident in the comparison of data from document

analysis and interviews. The interviewees (6 out of 36 graduating students) who admitted having written graffiti stated that they had written a name, a sign (doodling), a communication message, or an inscription about romance in class (publicly) but not politics/religion or sexuality, which comprised the main content of the graffiti found in the restrooms (i.e., non-public and purely anonymous). This anonymous nature of bathroom graffiti might lead us to conclude that bathroom graffiti is not suitable to be inquired about through follow-up interviews.

Considering the act of writing on public property as vandalism, the findings of the study can be interpreted as disappointing, since future teachers cannot be associated with such inappropriate behavior. Yet, if we consider graffiti as a part of the socialization process of adolescents and young adults, the apparent inconsistency between destructive behavior and the desirable personal qualities of teachers can encourage us to examine the unmet needs of future teachers in the formal and hidden teacher-training curricula. To illustrate, bathroom graffiti written by prospective teachers may point to the importance for student teachers to express themselves especially in terms of sexual, political, and religious issues and to socialize through proper communication with other students and faculty members. Although this may have been achieved to some extent, further progress can be obtained through curricular and/or extracurricular activities that provide prospective teachers with opportunities to communicate and interact socially on matters that concern them. Regular parties, open debates, discussion boards, Internet forums, and similar platforms for socialization are some examples of activities to prevent this destructive activity.

As a result of the analysis, there seem to be two main spheres where graffiti was produced with different topics and intentions: labs/classes and restrooms. The more socially acceptable ones, which were produced in labs/classes, focused on the need for belongingness, homesickness, romance, humor, or on the form of someone's name and signs (doodling). As suggested by the analysis of the interviews, a possible reason for students' attraction toward this activity was their instructors' failure to keep using stimulating pedagogies during class. Yet, when it comes to the students' choice of what graffiti to write, the question should be asked: Can the content of these inscriptions really imply some unmet needs? To what extent should we attribute the existence of many city/hometown names to the failure of the faculty to create a sense of belongingness or the existence of graffiti about romance to inadequate/unhealthy social relations between the sexes? These are very difficult questions to answer. An interpretation that is closer to the data, especially of the interviews, is the apparent failure by the faculty to engage students actively.

While these are the inferences based on the findings from the analysis of the public graffiti (i.e., in classes/labs), the more anonymous inscriptions can lead us to appreciate better the attitudes and feelings that prospective teachers hesitate to utter publicly. Consistent with the literature, sex seems to be the foremost taboo domain, disclosed mostly in private by men and women. Putting a side the sexual context, the content of other graffiti, such as politics/religion and humor, can be seen as a reflection of the cultural environment in which students live. The most interesting

point seems to be that men find more to say about politics/religion—even about matters that mostly concern women (e.g., what to wear). This finding seems to support the socio-politic inference as mentioned by Green (2003), which suggests that women deprived of liberal perspectives either refrain from thinking about such issues or are not even aware of their existential rights. This can be a result of Turkey's patriarchal culture in which women's behavior is something about which men have a say. Similarly this "miniature discussion board" reflects the dominance of men in discussing and deciding matters pertaining to religion/politics. In this male-dominant, conservative atmosphere, men might feel freer to write about humor and philosophy, whereas women may not even write about romance, unlike their counterparts in other cultures.

As a result of this study, some recommendations can be made based on the above findings and discussion to better link the counseling profession to teacher training. First, counseling should be employed as a major course in teacher education, highlighting the need to cultivate prospective teachers who can think without radical political/religious biases, judge and solve problems through respectful communication and negotiation, and welcome culturally diverse people and their worldviews (The Council for Accreditation of Counseling and Related Educational Programs [CACREP], 2009).

Second, it was clear from the findings that one reason for graffiti writing in classes or labs was boredom. Thus, faculty should improve their instructional pedagogies to activate student interest and involve students more in learning.

Finally, people have needs at all developmental levels in all multicultural contexts caused by a large spectrum of factors, including transitions across the lifespan as well as various situational events and environmental conditions (CACREP, 2009). The need for belongingness or feelings of homesickness may be two of these needs. In order to meet these needs in their students, faculty should welcome their students with an emotionally nurturing atmosphere and maintain such an atmosphere throughout the students' entire education. This can be achieved by organizing special programs and events, which aim at strengthening the solidarity and affection between prospective teachers and faculty.

References

- Aksu, M., Daloğlu, A., Yıldırım, S., Kiraz, E., & Demir, C.E. (2007). *Research on the profile of the students at faculties of education*. Ankara: METU.
- Bates, J. A., & Martin, M. (1980). The thematic content of graffiti as a nonreactive indicator of male and female attitudes. *Journal of Sex Research, 16*(4), 300-315.
- Beaton, D.E, Bombardier C., Guillemin F., & Ferraz, M.B. (2000). Guidelines for process of cross-cultural adaptation of self-report measures. *Spine, 25*, 3186-3191.

- Biklen, S.K., & Casella, R. (2007). *A practical guide to the qualitative dissertations*. New York&London: Teachers College Press.
- Bruner, E.M., & Kelso, J.P. (1980). Gender differences in graffiti: A semiotic perspective. *Women's Studies International Quarterly*, 3(2/3), 239-252.
- Center for Mental Health in Schools at UCLA. (2007). *Youth gangs and schools*. Los Angeles, CA: Author.
- The Council for Accreditation of Counseling and Related Educational Programs (CACREP) (2009). *2009 standards*. Retrieved March 28, 2009, from <http://www.cacrep.org/2009standards.pdf>
- Downe-Wamboldt, B. (1992). Content analysis: Method, applications, and issues. *Health Care for Women International*, 13, 313-321
- Gach, V. (1973). Graffiti. *College English*, 35(3), 285-287.
- Gadsby, J. (1995). *Looking at writing on the wall: A critical review and taxonomy of graffiti texts*. Retrieved July 05, 2008, from http://www.graffiti.org/faq/critical_review.html
- Green, J. (2003). The writing on the wall: Gender and graffiti. *Journal of Language and Social Psychology*, 22(282).
- Kan, K-H. (2001). Adolescents and graffiti. *Art Education*, 54(1), 18-23.
- Klingman, A., Shalev, R., & Peariman, A. (2000). Graffiti: A creative means of youth coping with collective trauma. *The Arts in Psychotherapy*, 27, 299-307.
- Knight, B.D. (1982). Identity and territory: Geographical perspectives on nationalism and regionalism. *Annals of the Association of American Geographers*, 72(4), 514-531.
- Kurt, İ. (2002). Gençlere rehber olmak açısından sraüstü yazılarının psikolojik analizi. İhsan Kurt (Ed.), *Psikolojiden Kültüre* (2nd ed.) (pp.229-275). Konya: Eğitim Kitapevi.
- Kuş, E. (2007). *Sosyal bilimlerde bilgisayar destekli nitel veri analizi* [Computerized qualitative data analyses in social sciences]. Ankara: Anı Yayıncılık.
- The Office of Juvenile Justice and Delinquency Prevention in the U.S. Department of Justice (1998). Wipe out vandalism and graffiti. *Youth in Action*. (Eric Document Reproduction Service No:434983)
- Richardson, D. (1990). *Is graffiti art? Proceedings of the In SEA 30th World Congress "Cultures and Transitions"*, 1990, Australian Institute of Art Education. Brisbane, Australia.
- Sechrest, L., & Olson, K. (1971) Graffiti in four types of institutions of higher education. *Journal of Sex Research*, 7(1), 62-71.

- Smith, M. (2003). Exploring target attractiveness in vandalism: An experimental approach. *Crime Prevention Studies*, 16, 197-236.
- Stocker, T.L., Dutcher, L.W., Hargrove, S.M., & Cook, E.A. (1972). Social analysis of graffiti. *Journal of American Folklore*, 85(338), 356-366.
- Walsh, M. (1996). *Graffiti*. Berkeley: North Atlantic Books.
- Yıldırım, A., & Şimşek, H. (2003). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri* [Qualitative Research Methods in Social Sciences]. Ankara: Seçkin Yayıncılık.

Öğretmen Eğitiminde Duvar Yazıları Üzerine bir Çalışma (Özet)

Problem Durumu: Duvar yazıları kişinin kendisini ifade etme yollarından birisidir. Özellikle gençler kendilerini dinleyecek birilerini bulamadıklarında, herhangi bir kamu malının üzerine yazılar yazmak suretiyle kendilerini ifade edebilirler. Hatta kendilerini dinleyecek birilerini bulsalar bile, özel ve yoğun hislerini başkalarına anlatmaktan çekinip, kimliklerini ifşa etmeyecek şekilde tuvalet, sıra, duvar gibi kamu mallarının üzerine yazılar yazmayı tercih edebilirler. Bu yüzden duvar yazıları psikolojik rehberlik ve danışmanlık alanının bir konusu olarak ele alınabilir. Bu çalışmada olduğu gibi duvar yazıları yoluyla duygu, düşünce ya da isteklerini ifade etmeye çalışan kişilerin öğretmen adayları olması halinde, duvar/sıra yazılarının incelenmesi, öğretmen adaylarını daha iyi anlamak suretiyle daha iyi öğretmen yetiştirme programları geliştirmek adına önemli bir fırsat olarak görülebilir. Bu yüzden, bu çalışmada öğretmen adayları tarafından yazılan duvar/sıra yazıları, rehberlik ve psikolojik danışma ile öğretmen yetiştirme alanlarını birleştiren disiplinler arası bir konu olarak ele alınmıştır.

Araştırmanın Amacı: Bu çalışmanın amacı öğretmen adayları tarafından yazılan duvar/sıra yazılarının içeriğini ve öğretmen adaylarının bu yazıları yazma nedenlerini incelemektir.

Yöntem: Bu çalışmada öğretmen adayları tarafından yazılan duvar/sıra yazıları hakkında derinlemesine bir inceleme yapabilmek amacıyla nitel içerik analizi ve yapılandırılmış nitel görüşme yöntemleri kullanılmıştır. Bu amaçla eğitim fakültesi binasının lavaboları ve derslik/laboratuvarlarından elde edilen toplam 178 yazı incelenmiştir. Ayrıca altı öğretmen adayı ile nitel görüşme yapılmıştır. Görüşme yapılan altı erkek öğretmen adayı toplam 36 (11 kadın ve 25 erkek) kişi arasından belirlenmiştir. Eğitim fakültesinden mezun olmak için gerekli ilişik kesme işlemlerini yürüten bu 36 son sınıf öğrencisi araştırma hakkında bilgilendirildikten sonra, kendilerine eğitimleri esnasında herhangi bir şekilde duvar/sıra yazısı yazıp yazmadıkları sorulmuş ve sadece altı erkek öğretmen adayı en az bir kez yazdığını ifade etmiştir. Devamında bu öğretmen adaylarına ne yazdıkları, nereye yazdıkları ve neden

yazdıklarıyla ilgili üç soru sorulmuştur. Nitel içerik analizi ve görüşmelerden elde edilen veriler Nvivo 8 nitel veri analizi yazılımı kullanılarak kodlanmış ve kategorilere ayrılmıştır.

Araştırmanın Bulguları: Yapılan analizler sonucunda öğretmen adayları tarafından yazılan yazıların sınıf ve laboratuvarlarda daha fazla (%71.9), lavabolarda ise daha az (%28.1) olduğu görülmüştür. Sınıf/laboratuvar yazıları, il veya ilçe isimlerinden (%25), kişi isimlerinden (%17.9), işaretlerden (karalama) (%16.4) ve romantizm içerikli ifadelerden (%10.9) oluşmaktadır. Erkek lavabolarındaki yazıların cinsellik konulu (%32.3), iletişim biçiminde (%29.4), politik/dini mesajlı (%23.5) ve mizah içerikli (%14.7) olduğu görülmüş, bununla birlikte bayan lavabolarında, sayı olarak erkeklerinden daha az olmakla birlikte, cinsellik ağırlıklı (%37.5) ve iletişim biçiminde (%37.5) duvar yazılarının olduğu saptanmıştır. Ayrıca diğer kültürlerde yapılan çalışmalarındaki bulgulara kıyasla bayan lavabolarında romantizmle ilgili nispeten az sayıda duvar yazısına rastlanmıştır. Araştırmada ayrıca duvar/sıra yazılarıyla ilgili çalışmaların yöntemsel sıkıntılarının varlığı saptanmıştır.

Araştırmanın Sonuçları ve Önerileri: Her ne kadar kamusal bir binada duvarlara/sıralara yazı yazma eylemi bir tür vandalizm olarak değerlendirilse de, özellikle lavabolara yazılmış olan yazıların, cinsellik, siyaset ve din konularında öğretmen adaylarının kendilerini ifade etme ve diğer öğrenci ve öğretmenlerle iletişim kurarak sosyalleşme ihtiyaçlarına işaret ettiği söylenebilir. Öğretmen adaylarının bu yöndeki ihtiyaçlarının daha fazla karşılanabilmesi amacıyla kendilerini farklı alanlarda özgürce ifade edebilecekleri ve düzeyli bir sosyal iletişim kurabilecekleri program içi ve program dışı fırsatların yaratılması gerekmektedir.

Ayrıca cinsellik, siyaset/din ve mizah içerikli yazıların kadın lavabolarından çok erkek lavabolarında bulunması, özellikle ilgili alan yazındaki bulguların aksine romantizmle ilgili yazılara kadın lavabolarında rastlanamaması da çalışmanın manidar bulgularından birisi olarak görülmüştür. Bu durum, ilgili alan yazındaki tespitler de dikkate alınarak, söz konusu kadın örneklem grubunun liberal değil, muhafazakâr bir bakış açısına sahip olmasıyla ya da içerisinde yaşadıkları ve etkilendikleri toplumun bu baskın özelliklere sahip olmasıyla açıklanabilir. Eğitim fakültesindeki öğrenimleri esnasında en az bir kez duvar/sıra yazısı yazdığını belirttiği için nitel görüşme yapılan öğretmen adaylarının tamamının erkek olması bu yorumu güçlendirmektedir. Bir başka deyişle, muhafazakâr toplumlarda kadınların duvar yazısı yazma oranı düşüktür. Diğer taraftan sınıf/laboratuvarlardaki sıra veya duvar yazılarının ders esnasında yazıldığı varsayıldığında, bunun temel nedeninin öğretim elemanının dersi öğrenciler için yeterince ilgi çekici hale getirememesi olduğu söylenebilir. Bu çıkarımın, nitel görüşmelerden elde edilen verilerle desteklendiği görülmektedir. Ancak bu şekilde derse aktif olarak katılmayan ve ders esnasında sıkılan öğrencilerin neden özellikle belirli kategorilerine giren içerikte yazılar yazdıklarını açıklamak oldukça güçtür. Buna göre yoğun bir şekilde yazılmış olan şehir/ilçe ve spor takımı isimlerinin aidiyet ihtiyacı veya ev/aile özlemi ile bağlantılı olabileceği yorumu yapılabilir. Son olarak, gerek bir vandalizm eylemi olduğu için

gerekse bazı yazıların içeriğinin müstehcen veya siyasi olmasından ötürü, söz konusu yazıları yazan kişilerin doğal olarak isimlerini gizli tutmaya çalıştıkları görülmüştür. Dolayısıyla bu ve benzeri çalışmalarda başta görüşmelerden olmak üzere elde edilen verilerin tam ve doğru olmama riski vardır. Bu yöntemsel sıkıntı incelenen dokümanların gerçekliği için de geçerlidir. Örneğin duvar/sıra yazıları farklı bir fakülteden gelen bir öğrenci tarafından ya da kadın lavabosundaki bir yazı bir erkek tarafından yazılmış olabilir. Bu ve benzeri etmenleri kontrol etmek bu tür çalışmalarda oldukça güçtür.

An Investigation of University Students' Critical Thinking Disposition and Perceived Problem Solving Skills*

Songül Tümkaya**

Birsel Aybek***

Habibe Aldağ****

Suggested Citation:

Tümkaya S., Aybek, B., & Aldağ, H. (2009). An investigation of university students' critical thinking disposition and perceived problem solving skills. *Eğitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 57-74.

Abstract

Problem Statement: Research studies related to developing critical thinking and problem solving as cognitive thinking skills are increasing in Turkey; however, the affective side of the issue is still being neglected. Critical thinking dispositions and perceived problem solving skills are important factors for determining educational needs of students.

Purpose of Study: The main purpose of this study was to determine the differences in university students' critical thinking disposition and perceived problem solving skills based on gender, grade level, and field of study. The relationship between critical thinking disposition and problem solving skills was also examined.

Methods: This research was conducted at a university setting in Turkey. Participants consisted of 353 volunteer students from various departments. Facione and Facione's (1996) "California Critical Thinking Disposition Inventory" and Heppner and Petersen's (1982) "Problem Solving Inventory" were used to gather the data. A Personal Information Form was developed to collect data on students' demographics. A Pearson product-moment correlations and multivariate test (MANOVA) was used with a .05 significance level to analyze the data.

* This article was presented at Affective Education in Action. The International Conference at the Çukurova University Adana, Turkey, June 28-30, 2007.

** Assoc. Prof. Dr., Cukurova University, Faculty of Education, Department of Primary Education, Adana, Turkey. stumkaya@cu.edu.tr

*** Lecturer Dr., Cukurova University, Faculty of Education, Department of Educational Sciences, Adana, Turkey. baybek@cu.edu.tr

**** Assist. Prof. Dr., Cukurova University, Faculty of Education, Department of Computer Education and Instructional Technology, Adana, Turkey. habibe@cu.edu.tr

Findings and Results: The result of this study indicates that a better disposition towards critical thinking was associated with greater problem solving skills. We found that gender was not a significant variable related to critical thinking disposition or to perceived problem solving skills. The students' field of study was related to critical thinking dispositions but it was not related to perceived problem solving skills; social science students had higher levels of these variables compared to science students. Grade level was significantly related to both problem solving and critical thinking dispositions. This result indicates that the students' levels of problem solving skills and critical thinking dispositions increase as they continue in their college education. Instructional programs for supporting critical thinking and problem solving skill, without neglecting the affective side of these concepts, must be incorporated into college curricula.

Conclusions: Problem solving skills are related to critical thinking dispositions. A university education is positively correlated with perceived problem solving skills as well as critical thinking dispositions. Field of study is another factor that determines critical thinking disposition levels. In addition, the interaction between gender and field of study was positively related to problem solving skills.

Recommendations: Increasing activities which develop critical thinking and problem solving skills through counseling services and integrating these activities into the instructional program would help students in their development of these skills.

Keywords: California critical thinking disposition inventory, critical thinking, problem solving inventory, university students

Ever since John Dewey argued, in 1933, that learning to think and reason were fundamental goals of education, critical thinking and problem solving have received overwhelming support from academicians. However, critical thinking and problem solving are still recurrent concerns for educators (Barnett, 1997; Halpern, 1998; McBride, Xiang, & Wittenburg, 2002). Despite the abundance of research related to critical thinking dispositions and perceived problem solving skills, little attention has been given to how they might differ among cultural groups. Within the limited literature, Brown (1977) has suggested that modes of thought are embedded in cultural milieu. A series of studies conducted with university students from mainland China, Hong Kong, and the United States of America have revealed that thinking styles vary across cultures (Zhang & Sternberg, 2001). For example, East Asians are more field dependent than are individuals in the USA (Witkin & Berry, 1975); East Asians tend to see social behavior as the outcome of a person-context interaction in contrast to Americans who tend to view social behavior as an unfold disposition (Choi, Nisbett, & Norenzayan, 1999). Therefore, the issue of cultural variability in thinking styles should not be ignored and more research is needed in this area. Accordingly, we aim to provide information on critical thinking dispositions and perceived problem solving abilities among Turkish individuals. The

purpose of this study is to determine Turkish university students' disposition of critical thinking and perceived problem solving skills based on gender, grade level, and field of study. We briefly present the definitions of the variables in this study and discriminate the differences between the concept of ability and disposition so that our readers can understand how we approach the study variables.

Critical Thinking

The definition of critical thinking varies because of its abstract characteristics. There is no universal agreement upon a definition for critical thinking (Brookfield, 1987; Thurmond, 2001; Yeh, 2002), and therefore a wide range of views exist as to what critical thinking is. For example, critical thinking has been conceptualized as a generic skill (Ennis, 1987; Halpern, 1998); an embedded skill (McPeck, 1981); deductive and inductive reasoning (Ennis, 1995); problem solving (Watson & Glaser, 1980); a subcomponent of lifelong learning skills (Candy, Crebert, & O'leary, 1994); reflective skepticism (McPeck, 1981); challenging assumptions (Brookfield, 1987); disciplined and dialectical thinking (Paul, 1990); purposeful and self-regulatory judgment (Facione & Facione, 1996); and critical being (Barnett, 1997).

While much attention focuses on the discrete cognitive skills associated with critical thinking, in the past two decades an interest in the affective or dispositional side of critical thinking has emerged (e.g., Baron, 1985; Perkins, Jay, & Tishman, 1993). Watson and Glaser (1980) incorporated the three important dimensions of attitude, knowledge, and skills necessary to engage in rational reasoning as a part of their definition of what a critical thinker is. Perkins et al. (1993) present a triadic dispositional theory, suggesting that intellectual behavior requires three logically separable components: ability, sensitivity, and inclination. Similarly, Facione, Facione, and Giancarlo (1998) described the concept of disposition as the consistent internal motivation to engage in problems and make decisions by using thinking. Dispositions are natural tendencies or habits of responding to particular situations in certain ways. These dispositions or tendencies to behave include a tolerance for ambiguity, a willingness to suspend judgment, and being open-minded, truthful, and analytical; in short, a willingness to engage in critical thinking (Beyer, 1987; Ennis, 1987, 1994, 1995).

According to Facione and Facione (1996), critical thinking is a nonlinear and cyclical process that allows people to make decisions on what to believe and what to do within a given context. We define critical thinking as the process of purposeful, self-regulatory judgment, which drives problem solving and decision making. Based on this definition, we consider, first, that this process requires knowledge to think critically to solve a problem. Second, it requires cognitive abilities such as evaluating the premises and thinking deductively or inductively. Third, it requires inclination and willingness to engage in critical thinking to solve problems. The disposition to think critically on a controversial issue might lead to a greater engagement in argumentation and, thus, to the development of cognitive skills or abilities. In

addition, education for critical thinking or problem solving requires attention to the cultivation of positive thinking dispositions and effective thinking skills. Therefore, we consider a possible relationship between perceived problem solving abilities and critical thinking dispositions.

Problem Solving

Teaching problem solving to improve students' thinking has been supported by several authors (Heppner & Petersen, 1982). The early focus, in the 1950s, was on basic science and these authors examined problem solving in various conditions such as solving water jar or string problems (Mayer, 1991). Emphasizing the importance of problem solving, it is suggested that education should focus on putting students into situations that simulate real scenarios so that they have opportunities to learn how to approach problems, formulate thinking, participate in discussions, and solve problems (Yeh, 2002). The focus of research shifted in the 1980s to how individuals appraise their abilities (Bandura, 1982). For instance, Butler and Meichenbaum (1981) hypothesized that if an individual's appraisal of his or her problem solving skills affects their problem solving performance, then it might also be an important predictor variable of the problem solving process.

Many authors also believe that dispositions toward critical thinking and cognitive skills for critical thinking must be taught and nurtured as early as possible in order to produce skillful problem solvers (Altun, 2003; Yeh, 2002). Problem solving is defined as a response given in an important and difficult situation where critical thinking is required for a solution (Altun, 2003). If critical thinking is a base for successful problem solving, then we might expect a relationship between the critical thinking dispositions and problem solving perceptions of an individual. However, critical thinking and problem solving are mostly investigated as independent from each other.

The Relation of Socio-Demographic Factors to Critical Thinking and Problem Solving

Gender is a demographic variable that has been examined within the context of perceived problem solving skills and a critical thinking. For example, Şahin, Şahin, and Heppner (1993) investigated the difference in perceived problem solving skills among American and Turkish university students and found that male American students were more confident than male Turkish students. They also found that American female students have more control over their problem solving skills than Turkish female students. Taylan (1990) studied gender differences in problem solving skills. He concluded that there is no gender difference related to problem solving skills. Gökçakan and Nadir (2001) studied problem solving skills as a predictive variable for self-image. They, too, found no gender differences. Dikici, Yavuzer, and Demir (2001) investigated the adaptation and problem solving skills of university students and gender was also an insignificant variable in this study. However, Ferah (2000) found a significant difference between male and female military-school students; female students showed a more positive and systematic

approach to problem solving compared to male students. Gender has also been investigated in the context of critical thinking dispositions. Kökdemir (2003) found that female students have a higher disposition for critical thinking compared to male students. Genç (2008) studied critical thinking tendencies among teachers. The results indicated that tendencies differ in sub-scales on open-mindedness and curiosity with regard to gender. However, Özdemir (2005) found that there is no significant difference in critical thinking dispositions between male and female students. The findings in Kawashima and Shiomi's study (2007) also revealed that there was no difference.

Grade level is another variable that has been analyzed in a limited number of studies on the Turkish population. Tümkaya and İflazoğlu (2000) investigated automatic thinking and problem solving skills based on grade level, gender, parents' education level, and parents' occupation and socioeconomic level among a group of college students. Grade level was found to be the only significant variable related to perceived problem solving skills. Moreover, Alver (2005) studied perceived problem solving skills with a university population. The researcher stated that grade level and major were the factors correlated with the problem solving skills. Dikici et al. (2001) found a decline in second and third year university students' problem solving skills and an increase in fourth year students. Öztürk and Ulusoy (2008) investigated the relationship between demographic factors and critical thinking dispositions in nursing students. The study group consisted of 22 graduate and 312 undergraduate nursing students at a large Turkish public college. The results showed that grade level, GPA, and age were positively related factors to critical dispositions. This conclusion is validated for the grade level factor in a similar population of nursing students in Turkey (Zaybak & Khorshid, 2006). Brooks and Shepherd (1992) studied the relationship between the level of professionalism and critical thinking in nursing students; they found that years of experience was an influencing factor on the level of professionalism and critical thinking. Kawashima and Shiomi (2007) found that seniors have higher critical thinking dispositions compared to freshman students. Aybek (2006) investigated the effect of subject-based and skill-based critical thinking instructional programs among student-teachers. In this experimental study, a group of senior student-teachers was defined as the skill-based experimental group, a group of sophomore student-teachers was defined as the subject-based experimental group, and a group of junior student-teachers was defined as the no-treatment control group. The researcher concluded that senior students' critical thinking level was much higher than the other groups. However, there were also confounding results about grade level as a related factor to critical thinking dispositions. Ip, Lee, Lee, Chau, Wotton, and Chang (2000) studied the critical thinking disposition of a group of nursing students in Hong Kong; they found that juniors had the highest levels of critical thinking dispositions, freshmen had the second highest levels, and sophomores had the lowest levels of critical thinking dispositions.

The gender factor has been well-defined in relationship to perceived problem solving skills in a limited number of studies conducted with Turkish samples. However, this factor is not yet clear when it comes to the critical thinking dispositions of the Turkish populations since critical thinking dispositions as well as perceived problem solving abilities are new topics for Turkey. Furthermore, these limited studies on critical dispositions in Turkey are conducted on a homogeneous population, such as students who major in mathematics education (Türnüklü & Yeşildere, 2005) or nursing students (Öztürk & Ulusoy, 2008; Zaybak & Khorshid, 2006). More research is needed to clarify the role of gender, as well as the field of study and grade level on the critical thinking dispositions and the perceived problem solving skills of Turkish university student population.

Research Questions:

Consequently, research questions guiding the study included:

1-What is the relationship between college students' critical thinking disposition and problem solving skills?

2-How do the levels of students' critical thinking disposition and problem solving skills differ based on the students' gender, grade level, and field of study?

Method

Participants

This research was conducted at a large state university in Turkey during the 2006-2007 academic year. A total of 353 students, 204 social science students (Education=61, Fine Arts=54, Theology=45, Economics and Administrative Sciences=4) and 149 science students (Fisheries=48, Engineering and Architecture=44, Sciences and Literature=57), volunteered to participate in the research. The sample was made up of 50% male and 50% female students. The mean age of participants was 22.30 ($SD=2.50$).

Measures

The California Critical Thinking Dispositions Inventory (CCTDI; Facione et al., 1998) and the Problem Solving Inventory (Heppner & Petersen, 1982) served as the data-gathering instruments for the study. The Participant Information Form was developed to collect data on participants' socio-demographic attributes.

The California Critical Thinking Dispositions Inventory (CCTDI). The CCTDI (Facione et al., 1998) addresses critical thinkers' affective and attitudinal aspects. The CCTDI contains 75 Likert-style items and is designed for use among college/university students. The items are answered using a 6-point Likert scale in which one reflects a strong agreement and six reflects a strong disagreement. In the CCTDI, critical thinking disposition can be measured by means of seven subscales: Truth-seeking, Open-mindedness, Analyticity, Systematicity, Inquisitiveness, Self-Confidence, and

Maturity (Facione et al., 1998). The total CCTDI score may range from 70 to 420. The final score reflects the strength of the participant's disposition to think critically. A score over 350 suggests a strong disposition, whereas a score below 280 is representative of a weak disposition to think critically. Scores between 280 and 350 indicate a positive inclination toward critical thinking (Facione, Facione, & Sanchez, 1994; Profetto-McGrath, Hesketh, Lang, & Estabrooks, 2003).

Although convergent validity studies between various critical thinking disposition measures are not available, Facione et al. (1994) reported significant correlations supporting the concurrent validity between individual CCTDI scales and psychological scales that targeted the observed constructs. Reliability tests reported an overall test score of .90 (Cronbach's alpha); sub-scale scores ranged from .71 to .80. Studies using the CCTDI consistently report a satisfactory reliability for the instrument and its subscales (Facione & Facione, 1996; Jacobs, 1995). Ip et al. (2000) reported the scale's internal consistency as .90 for the Chinese version.

The CCTDI was adapted into the Turkish language by Kökdemir (2003). The analysis of the data collected from a group of college students showed a similar structure to the original one; however, two factors were combined into one factor. Thus, the Turkish version of the CCTDI contains seven subscales and 51 items. In Kökdemir's study, the Cronbach alpha was .88. In another study, Aybek (2006) found that the test-retest reliability of the scale was .80. Similarly, the current study found the internal consistency of the scale to be .78. We also determined the test-retest reliability to be .74 for 60 college students.

The Problem Solving Inventory (PSI). Rather than to assess actual problem solving capabilities, the PSI (Heppner & Petersen, 1982; Heppner, Witty, & Dikson, 2004) was designed to assess an individual's perceptions or awareness of one's own problem solving behaviors and attitudes. This instrument was used to measure students' self-confidence and feelings of self-control when problem solving, as well as the way in which the individual approaches problem solving. PSI uses a 6-point Likert scale, ranging from 1 (*Strongly Agree*) to 6 (*Strongly Disagree*). The items contain positive and negative judgments about problem solving. The negative judgments are reversed for evaluation. Thus, a low score of PSI indicates a positive attitude and perception of effective problem solving abilities. In contrast, high scores indicate an inability to reach a successful solution when faced with a problem. In other words, a high score of PSI means that the person perceives him or herself incapable of problem solving. The total number of items ($N=35$) is divided into three subscales: Problem solving Confidence, Approaching-Avoidance Style, and Personal Control. The Cronbach alpha coefficient for the total inventory was .89 (Heppner & Petersen, 1982). The PSI has been found to have acceptable internal consistency estimates across a number of populations and cultures (e.g., Heppner & Wang, 2003). Summing across studies, the PSI total obtains average alpha coefficients in the high .80s.

The inventory's reliability and validity for the Turkish culture has been tested by Taylan (1990), by Şahin, Şahin, and Heppner (1993), and by Çam (1995). Statistical tests of the instrument's validity indicated that the measurement model was highly valid. Taylan (1990) and Çam (1995) found a high relation among subscales which resulted in a factor loading different than that of the original form. For this reason, they strongly suggested that the total score of the PSI must be used instead of using subscale scores separately. The Turkish version of the Problem Solving Inventory contains 32 items out of 35. High scores indicate ineffective problem solving skills as in the original version.

In this current study, the item-item correlation ranged from .33 to .66, the Cronbach alpha for the total scale was .87, and the test-retest reliability was .83. We also decided to use the total PSI score as a result a high relation among subscales, adapting to Taylan (1990) and Çam's (1995) suggestions.

Participant Information Form. This form was developed by researchers to collect data on participants' socio-demographic attributes such as age, gender, grade level, and field of study.

Data Analysis

A Pearson product-moment correlation analysis and a multivariate test ($2 \times 2 \times 2$ MANOVA) were used to analyze the data. SPSS 11.05 was used for the statistical data analysis; a significance level of .05 was adopted as the level of significance.

Findings

To answer the first research question, the Pearson product-moment correlation analysis was used. A significant negative correlation was found between participants' level of critical thinking disposition and perceived problem solving skills ($r = -.13, p < .05$). It is important to note that low scores are representative of a positive judgment of one's problem solving abilities and high scores indicate a poorer judgment of problem solving abilities. This finding indicates that a higher disposition towards critical thinking was associated with greater problem solving skills.

A multivariate analysis of variance was conducted to determine if gender, grade level, and field of study variables had a significant correlation with students' critical thinking disposition and problem solving skills. The means and standard deviations for the CCTDI and PSI scores is shown in Table 1. Equality of covariance was acceptable based on Box's M (Box's M: 19.655; $F = .918, p > 0.05$). The result of the Levene's test revealed an equality of variance on dependent variables ($F(7, 345) = 1.345, p > .05$; $F(7, 345) = .827, p > .05$).

Table 1***Means and Standard Deviations for Problem Solving and Critical Thinking Disposition related to Grade Level, Gender, and Field of Study Factors***

Variables	Grade Level	Gender	Field of Study	M	SD	N
<i>Problem Solving*</i>	Freshman	Women	Social Sciences	84.34	19.41	59
			Science	92.83	23.66	35
		Men	Social Sciences	93.28	19.91	42
			Science	93.16	19.71	43
	Senior	Women	Social Sciences	79.62	21.99	45
			Science	86.11	23.28	36
		Men	Social Sciences	87.91	24.27	58
			Science	78.26	19.56	35
	Total	Women	Social Sciences	82.30	20.60	104
			Science	89.42	23.55	71
		Men	Social Sciences	90.17	22.59	100
			Science	86.47	20.89	78
<i>Critical Thinking Disposition</i>	Freshman	Women	Social Sciences	192.64	16.37	59
			Science	186.77	15.44	35
		Men	Social Sciences	187.07	15.62	42
			Science	191.23	12.99	43
	Senior	Women	Social Sciences	212.82	15.87	45
			Science	205.86	12.35	36
		Men	Social Sciences	212.14	13.69	58
			Science	206.68	13.86	35
	Total	Women	Social Sciences	201.37	18.96	104
			Science	196.45	16.87	71
		Men	Social Sciences	201.61	19.07	100
			Science	198.17	15.37	78

* Lower score of PSI refers to higher problem solving skills

Among female students, the senior students in the social sciences had the lowest mean scores on the PSI ($M=79.62$). Among male students, the senior students in science had the lowest mean scores on the PSI ($M=78.26$). The total PSI scores of female students in social science were lower than the total PSI scores of female students in science ($M=82.30$). Senior female and senior male students in social science had the highest mean score on the CCTDI ($M(W)=212.82$; $M(M)=212.14$). The total mean score of male students in social science on the CCTDI ($M=201.61$) was higher compared to the total mean score of male students in science ($M=198.17$). The results of the MANOVA are given in Table 2.

Table 2

The Result of the Multivariate Analysis of Variance for Problem Solving and Critical Thinking Disposition related to Grade Level, Gender, and Field of Study Factors

Source	Dependent Variables	df	F	P	η^2
Gender (A)	Problem Solving	1	1.079	.300	.003
	Critical Thinking	1	.023	.879	.000
Grade Level (B)	Problem Solving	1	11.487	.001	.032
	Critical Thinking	1	156.719	.000	.312
Field of Study (C)	Problem Solving	1	.309	.579	.001
	Critical Thinking	1	4.912	.027	.014
A x B	Problem Solving	1	.893	.345	.003
	Critical Thinking	1	.039	.844	.000
B x C	Problem Solving	1	1.520	.219	.004
	Critical Thinking	1	2.819	.094	.008
A x C	Problem Solving	1	7.002	.009	.020
	Critical Thinking	1	3.280	.071	.009
A x B x C	Problem Solving	1	.648	.421	.002
	Critical Thinking	1	1.789	.182	.005

The analysis indicated that there was no significant difference among groups based on gender; however, grade level was significantly correlated with problem solving ($F(1,345)=11.487$; $p<.05$) and critical thinking ($F(1,345)=0.023$, $p>.05$) scores (Wilk's Lambda (Λ)=0.676; $F(2,344)=82.370$, $p<.05$). Mean scores showed that senior students' problem solving scores were lower than the freshman students' problem solving scores. These results indicated that seniors were more skilled problem solvers compared to freshmen. The results also revealed a significant difference in critical thinking skills in favor of senior students. Based on these results, we can conclude that an increase in years spent in school is associated with higher levels of critical thinking disposition and higher levels of perceived problem solving skills.

Field of study was another variable that was significantly related to the dependent variables (Wilk's Lambda (Λ)=0.985; $F(2,344)=2.560$, $p<.05$). Critical thinking scores were differentiated based on field of study ($F(1,345)=4.912$; $p<.05$); however, no significant differences were found among problem solving scores. Social science students' critical thinking scores were higher than the science students' scores.

Other results indicated that the interaction between gender and grade level and the interaction between grade level and the field of study were not significant. However, the interaction between gender and field of study had a significant relationship to the dependent variables (Wilk's Lambda (Λ)=.972; $F(2,344)=4.951$, $p<.05$). Based on this interaction effect, there is a significant difference between problem solving scores; however, the interaction was not significant for critical thinking scores. The mean scores on problem solving indicated that female students in social science had higher perceived problem solving skills than female students in

science. Conversely, the mean scores of problem solving indicated that male students in science had higher perceived problem solving skills than male students in social science. Finally, the interaction between gender, grade level, and field of study was not significantly related to problem solving or critical thinking scores.

Discussion and Recommendations

The results of our study indicated that a better disposition towards critical thinking is associated with greater problem solving skills. This finding is in line with Kökdemir's (2003) study in which it was also found that students who had higher levels of critical thinking skills also had higher levels of problem solving and decision-making skills.

The results in our study show that there is not a significant difference either in problem solving skills or in critical thinking dispositions based on gender. This result is in line with previous research which found gender to be an insignificant variable in terms of university students' critical thinking dispositions (Kawashima & Shiomi, 2007; Özdemir, 2005; Taylan, 1990) and problem solving skills (Alver, 2005; Dikici et al. 2001; Gökçakan & Nadir, 2001; Taylan, 1990; Tümkaya & İflazoğlu, 2000). However, our results contradict the finding of a related study by Ferah (2000) that indicated that female students have a more positive and systematic approach to problem solving compared to male students.

Another result of our study was that there were significant differences in perceived problem solving skills and critical thinking dispositions based on grade level. Senior students had higher levels of critical thinking disposition. These results are supported by similar studies (Aybek, 2006; Brooks & Shepherd, 1992; Kawashima & Shiomi, 2007; Öztürk & Ulusoy, 2008; Zaybak & Khorshid, 2006). Senior students also had higher levels of problem solving skills. This result similarly is supported by previous research (Alver, 2005; Dikici et al., 2001; Tümkaya & İflazoğlu, 2000). This finding indicates that students' levels of perceived problem solving skills and critical thinking disposition increase as they progress through their college education. Nonetheless, the results of the Ip et al. study (2000) stand in contrast with our study's findings concerning critical thinking disposition. Ip et al. (2000) found that 3rd year nursing students scored significantly lower than their junior counterparts in other fields of study.

The field of study variable was significantly related to the students' critical thinking disposition, favoring social science students over science students, but not on perceived problem solving skills. This result contradicts with Özdemir's finding (2005) that there is no difference for critical thinking disposition. However, findings related to problem solving are partially in contrast with Alver's study (2005) results. Alver's study showed that social science students had better problem solving skills compared to science students. Considering the inconsistent results in these different studies, we emphasize the necessity of integrating new strategies and teaching methods that support not only problem solving, but also critical thinking skills into

all programs. Constructivist approaches, teamwork, and psychological and instructional interaction support might be highlighted.

Our study is limited to the 353 students attending seven different departments/schools. We combined departments/schools into a field variable: social science and science. Increasing the number of participants for each school and analyzing the data at the department level might give a clearer picture of differences among departments on critical thinking dispositions and perceived problem solving skills. This study contributes to literature by providing information about the relationship of critical thinking dispositions and perceived problem solving skills to gender, field of study, and grade level among Turkish university students. It also underlines the relationship between critical thinking dispositions and perceived problem solving skills. The results of this study also offer information for the redesigning of counseling services on college campuses and in instructional programmes. Although the significant difference among grade levels favors seniors and is a positive indicator of years spent in higher education, the total scores are lower than Kökdemir's (2003) total scores of critical thinking among the Turkish population. This indicates that interventional studies may be necessary for this particular population of students. Students who are low in critical thinking dispositions and perceived problem solving might be directed to the counseling services for individual or group intervention studies. Counseling services could provide students with activities related to solving daily life problems. Faculty members should be informed about the student's level of critical thinking and problem solving levels. Each department could provide opportunities for solving problems related to the discipline being studied. More research on the Turkish university population is needed to achieve a clear picture about critical thinking and problem solving in terms their relationship to demographic factors.

In summary, the results of this study indicate that a greater disposition towards critical thinking was associated with greater problem solving skills. We found that there were no gender differences in terms of critical thinking disposition and problem solving skills. Social science students had higher levels of critical thinking dispositions compared to science students. However, grade level was significantly related to the both variables, indicating the positive contribution of time spent in education. Instructional programs for supporting critical thinking and problem solving must be incorporated into the college curricula.

References

- Altun, İ (2003). The perceived problem solving ability and values of student nurses and midwives. *Nurse Education Today*, 23, 575-584.
- Alver, B. (2005). Üniversite öğrencilerinin problem çözme becerileri ve akademik başarılarının çeşitli değişkenlere göre incelenmesi [An investigation of university students problem solving skills and academic achievement based on various variables]. *VIII. Ulusal Psikolojik Danışma ve Rehberlik Kongresi Bildiri Özetleri Kitapçığı*, Marmara Üniversitesi, 21-23 Eylül, İstanbul, 188-189.

- Aybek, B. (2006). *Konu ve beceri temelli eleştirel düşünme öğretiminin öğretmen adaylarının eleştirel düşünme eğilimi ve düzeylerine etkisi* [The effect of content and skill based critical thinking teaching on prospective teachers' disposition and level in critical thinking]. Unpublished doctoral dissertation, Çukurova Üniversitesi, Sosyal Bilimler Enstitüsü, Adana.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122-147.
- Barnett, R. (1997). *Higher education: A critical business*. Buckingham, UK: Society for Research into Higher Education and Open University Press.
- Baron, J. (1985). *Rationality and intelligence*. New York: Cambridge University Press.
- Beyer, B. (1987). *Practical strategies for the teaching of thinking*. Boston: Allyn and Bacon.
- Brookfield, S.D. (1987). Developing critical thinkers. *Challenging adults to explore ways of thinking and acting*. San Francisco: Jossey-Bass.
- Brooks, K. L. & Shepherd, J.M. (1992). Professionalism versus general critical thinking abilities of senior nursing student in four types of nursing curricula. *Journal of Professional Nursing*, 8 (2), 87-95.
- Brown, A. L. (1977). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (pp.77-165). Hillsdale, NJ: Erlbaum.
- Butler, L. & Meichenbaum, D. (1981). The assessment of interpersonal problem solving skills. In P. C. Kendall & S. D. Hollen (Eds.), *Assessment strategies for cognitive behavioral interventions* (pp. 197-225). New York: Academic Press.
- Candy, P., Crebert, G., & O'leary, J. (1994). Developing lifelong learners through undergraduate education, *NBEET Commissioned Report*, No. 28, (Canberra, Australian Government Printer).
- Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Casual attributions across cultures: variation and universality. *Psychological Bulletin*, 125, 47-63.
- Çam, S. (1995). Öğretmen adaylarının ego durumları ile problem çözme becerisi algısı ilişkisinin incelenmesi [Relation of ego states and problem solving skills perception among teacher candidates]. *Psikolojik Danışma ve Rehberlik Dergisi*, 6(2), 37-42.
- Dewey, J. (1933). *How we think*. New York: Heath.
- Dikici, H., Yavuzer, Y., & Demir, Z. (2001). Üniversite öğrencilerinin uyum düzeyleri ile problem çözme yeteneklerinin çeşitli değişkenler açısından incelenmesi [An investigation of university students problem solving skills and level of adaptation based on various variables]. X.Ulusal Eğitim Bilimleri Kongresi, A. İ. Baysal Üniversitesi. 7-9 Haziran, Bolu.
- Ennis, R. (1987). A taxonomy of critical thinking dispositions and abilities, In J. Baron & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9-26), New York: W.H. Freeman.
- Ennis, R. (1994). *Critical thinking dispositions: theoretical and practical considerations in their delineation, endorsement, and assessment*. Unpublished manuscript, University of Illinois - Champaign.
- Ennis, R. H. (1995). Critical thinking dispositions: analysis and extension, putative visa, and assessability. *Inaugural symposium of the Eastern Shore Institute for the advancement of thinking*, Salisbury, MD: Salisbury State University.

- Ferah, D. (2000). *Kara harp okulu öğrencilerinin problem çözme becerilerini algulamalarının ve problem çözme yaklaşım biçimlerinin cinsiyet, sınıf, akademik başarı ve liderlik yapma açısından incelenmesi* [Investigation of perception of problem solving skills and problem solving approaches of the Turkish army academy cadets in terms of gender, grade, academic success and being leader]. Unpublished master dissertation, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Facione, P. A., Facione, N. C., & Giancarlo, C. A. F. (1998). *The California Critical Thinking Dispositions Inventory*. Millbrae, CA: California Academic Press.
- Facione, N, Facione, P., & Sanchez, M. (1994). Critical thinking disposition as a measure of competent clinical judgment: the development of the California critical thinking disposition inventory. *Journal of Nursing Education*, 33, 345-350.
- Facione, N. C. & Facione, P. A. (1996). Externalizing the critical thinking in knowledge development and clinical judgment. *Nursing Outlook*, 44 (3), 129-136.
- Genç, S. Z. (2008) Critical thinking tendencies among teacher candidates. *Educational Sciences: Theory and Practice*, 8(1), 107-116.
- Gökçakan, Z. & Nadir, B. (2001). Ergende problem çözme becerisini yordayıcı bir değişken olarak benlik imgesi [Self-perception as a predictive variable of problem solving skills on adolescent population]. *VI. Ulusal Psikolojik Danışma ve Rehberlik Kongresi Bildiri Özetleri Kitabı*, ODTÜ, 5-7 Eylül, Ankara, 17.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist*, 53, 449-455.
- Heppner, P. P. & Wang, Y. (2003). Problem solving appraisal. In S. J. Lopez & C. R. Snyder (Eds.), *Positive psychological assessment: A handbook of models and measures* (pp. 127- 138). Washington, DC: American Psychological Association.
- Heppner, P. P., Witty, T. E., & Dixon, W. A. (2004). Problem solving appraisal and human adjustment: A review of 20 years of research using the problem solving inventory. *The Counseling Psychologist*, 32(3), 344-428.
- Heppner, P. P. & Petersen, C. H. (1982). The development and implications of a personal problem solving inventory. *Journal of Counseling Psychology*, 29 (1), 66-75.
- Ip, W. Y., Lee, D. T., Lee, I. F., Chau, J. P., Wootton, Y. S., & Chang, A. M. (2000). Disposition towards critical thinking: A study of Chinese undergraduate nursing students. *Journal of Advanced Nursing*, 32(1), 84-90.
- Kawashima, N. & Shiomi, K. (2007). Factors of the thinking disposition of Japanese high school students, *Social Behavior and Personality*, 35(2), 187-194.
- Kökdemir, D. (2003). *Belirsizlik durumlarında karar verme ve problem çözme* [Decision making and problem solving under uncertainty]. Unpublished doctoral dissertation, Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Mayer, R. E. (1991). *Thinking, problem solving, cognition*. New York: W. H. Freeman.
- McBride, R. E., Xiang, P., & Wittenburg, D. (2002). Dispositions toward critical thinking: The preservice teacher's perspective. *Teachers and Teaching: Theory and Practice*, 8(1), 29-40.
- McPeck, J. E. (1981). *Critical thinking and education*. Oxford, UK: Martin Robertson.
- Paul, R. W. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world*. Rohnert Park, CA: Sonoma State University.

- Perkins, D., Jay, E., & Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking, *Merrill-Palmer Quarterly*, 39, 1-21.
- Profetto-McGrath, J., Hesketh, K. L., Lang, S., & Estabrooks, C. E. (2003). A study of critical thinking and research utilization among nurses. *Western Journal of Nursing Research*, 25(3), 322-337.
- Özdemir, S. M. (2005). Üniversite öğrencilerinin eleştirel düşünme becerilerinin çeşitli değişkenler açısından değerlendirilmesi [An evaluation of university students critical thinking skills with some variables]. *Türk Eğitim Bilimleri Dergisi*, 3(3), 297-314.
- Öztürk, N. & Ulusoy, H. (2008). Baccalaureate and masters' degree nursing students' levels of critical thinking and factors influencing critical thinking. *Maltepe Üniversitesi Hemşirelik Bilim ve Sanatı Dergisi*, 1(1), 15-25.
- Şahin, N., Şahin, N., & Heppner, P. P. (1993). Psychometric properties of the Problem Solving Inventory in a group of Turkish university students. *Cognitive Therapy and Research*, 17, 379-396.
- Taylan, S. (1990). *Heppner'in problem çözme envanteri'nin uyarlama, geçerlik ve güvenirlik çalışmaları* [The validity and reliability studies of the Heppner' problem solving inventory]. Unpublished master dissertation, Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Tümekaya, S. & İflazoğlu, A. (2000). Ç.Ü. sınıf öğretmenliği öğrencilerinin otomatik düşünce ve problem çözme düzeylerinin bazı sosyo-demografik değişkenlere göre incelenmesi [An investigation of automatic thought and problem solving levels of students in classroom teaching unit of Ç.U. according to some socio-demographic variables]. *Ç.Ü. Sosyal Bilimler Dergisi*, 6(6), 143-158.
- Thurmond V. A. (2001). The holism in critical thinking. *Journal of Holistic Nursing*, 19, 375-389.
- Türmüklü, E. B. & Yeşildere, S. (2005). A profile from Turkey: Critical thinking dispositions and abilities of pre-service mathematics Teachers of 11-13 year. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 38(2), 167-185
- Watson G. & Glaser E. M. (1980). *Watson-Glaser critical thinking appraisal forms A & B*. New York: The Psychological Corporation.
- Witkin, H. A. & Berry, J. W. (1975). Psychological differentiation in-cross cultural perspective. *Journal of Cross Cultural Psychology*, 6, 4-87.
- Yeh, M. (2002). Assessing the reliability and validity of the Chinese version of the California critical thinking disposition inventory. *International Journal of Nursing Studies*, 39, 123-132.
- Zaybak, A. & Khorshid, L. (2006). The investigation of the level of critical thinking of students in Ege University School of Nursing. *Ege Üniversitesi Hemşirelik Yüksek Okulu Dergisi*, 22(2), 137-146.
- Zhang L. & Sternberg R. J. (2001). Thinking styles across cultures: Their relationships with student learning. In R. J. Sternberg, & L. Zhang (Eds.), *In perspectives on thinking, learning, and cognitive styles* (pp. 197-226), Mahwah NJ: Erlbaum.

Üniversite Öğrencilerinin Eleştirel Düşünme Eğilimleri ve Problem Çözme Becerilerinin İncelenmesi

(Özet)

Problem Durumu: Günümüzde eğitimden beklenen en önemli görev, demokratik, yaratıcı, üretici, eleştirel ve çok yönlü düşünebilen, öğrenmeyi öğrenen, problem çözebilen, insanlara saygılı ve düşüncelere hoşgörü ile bakabilen sorumlu vatandaşlar yetiştirmektir. Birçok eğitimci eğitim kurumlarında ve özellikle üniversitelerdeki derslerin içeriğinin ancak düşünülerek ve problem çözerek öğrenilebileceğini aksi halde öğrencilerin birçok bilgiyi ezberleme yoluna gideceklerini belirtmektedir. Türkiye’de eleştirel düşünme ve problem çözme becerilerine ilişkin araştırma sayısında son yıllarda artış gözlenmektedir. Ancak diğer eğitimsel çözüm ve araştırmalarda olduğu gibi duyuşsal yön yine ihmal edilmektedir. Bu nedenle bu araştırmada üniversite öğrencilerinin eleştirel düşünme eğilimleri ile problem çözme becerileri arasındaki ilişkinin incelenmesi amaçlanmıştır.

Araştırmanın Amacı: Bu araştırmanın temel amacı, eleştirel düşünme eğilimi ile problem çözme becerisi arasında anlamlı bir ilişkinin olup olmadığını ortaya koymak ve ayrıca üniversite öğrencilerinin eleştirel düşünme eğilimleri ile problem çözme becerilerinin; cinsiyet, sınıf düzeyi ve eğitim gördükleri alanlara göre anlamlı bir şekilde farklılaşp farklılaşmadığını belirlemektir.

Araştırmanın Yöntemi: Araştırma Türkiye’de bir üniversitede, 2006-2007 bahar yarıyılında öğrenim gören öğrencilerle gerçekleştirilmiştir. Araştırmaya sosyal bilimler alanından 204, fen bilimleri alanından 149 gönüllü, öğrenci olmak üzere toplam 353 öğrenci katılmıştır. Araştırmaya katılan öğrencilerin % 50’si kız, % 50’si erkektir. Araştırmada veriler, Facione ve Facione’ın (1992) “California Eleştirel Düşünme Eğilimi Ölçeği” ve Heppner ve Petersen’in (1982) “Problem Çözme Becerisi” ölçeği ile toplanmıştır. Öğrencilerin sosyo-demografik özelliklerini (yaş, cinsiyet, sınıf düzeyi ve fakülte) belirlemek için ise araştırmacılar tarafından geliştirilen “Kişisel Bilgi Formu’ndan yararlanılmıştır. Verilerin analizinde Pearson momentler korelasyonu ve çok yönlü varyans analizi (MANOVA) kullanılmıştır. İstatistiksel işlemler SPSS 11.05 paket programı aracılığıyla yapılmıştır. Tüm analizlerde anlamlılık düzeyi .05 olarak kabul edilmiştir.

Araştırmanın Bulguları: Öğrencilerin eleştirel düşünme eğilimleri ile problem çözme becerileri arasında anlamlı bir ilişki olduğu görülmüştür. Analizler sonucunda öğrencilerin cinsiyetlerine göre bağımlı değişkenlerde farklılık olmadığı görülmüştür. Problem çözme ve eleştirel düşünme değişkenleri için elde edilen değerler anlamlı değildir.

Öğrencilerin sınıf düzeylerine göre problem çözme ($F(1,345) = 11.49; p < .05$) ve eleştirel düşünme ($F(1,345) = 156.72; p < .05$) toplam puanlarının ise farklılaştığı bulunmuştur (Wilk’s Lambda(Λ) = .676; $F(2,344) = 82.37, p < .05$).

Ortalamalara bakıldığında son sınıf öğrencilerinin, toplam problem çözme beceri puanlarının birinci sınıfta okuyan öğrencilerinkinden daha düşük olduğu görülmüştür. Problem çözme ölçeğinin özelliği nedeniyle (ortalamaların düşmesi problem çözme becerisinin arttığını gösterir) bu sonuç son sınıf öğrencilerinin, birinci sınıf öğrencilerinden daha yüksek bir problem çözme becerisine sahip olduklarını göstermektedir. Aynı şekilde toplam eleştirel düşünme puan ortalamalarına bakıldığında da son sınıf öğrencilerinin lehine bir durum olduğu anlaşılmaktadır. Bu sonuçlara bağlı olarak sınıf düzeyinin artmasıyla birlikte öğrencilerdeki problem çözme becerileri ile eleştirel düşünme eğilimlerinin arttığı söylenebilir.

Üniversite öğrencilerinin eğitim gördükleri alana göre bağımlı değişkenlerde farklılık olduğu görülmüştür (Wilk's Lambda(Λ)= .985; $F(2,344)=2.56$, $p < .05$). Eğitim gördükleri alana göre eleştirel düşünme puanları anlamlı bir şekilde farklılaşırken ($F(1,345) = 4.91$; $p < .05$), problem çözme beceri puanları arasında anlamlı bir farklılık bulunmamıştır. Ortalama değerleri sosyal bilimler alanında eğitim gören öğrencilerin eleştirel düşünme eğilimlerinin fen bilimlerinde okuyan öğrencilere göre daha yüksek olduğunu göstermiştir.

Cinsiyet ve sınıf düzeyinin ortak etkileşimi problem çözme ve eleştirel düşünme puanlarında anlamlı bir fark yaratmamıştır. Aynı şekilde sınıf düzeyi ile eğitim alanının ortak etkileşimi de problem çözme ve eleştirel düşünme puanlarında anlamlı bir fark yaratmamıştır.

Cinsiyet ve eğitim alanının ortak etkileşiminin bağımlı değişkenlerde ilişkili olduğu bulunmuştur (Wilk's Lambda(Λ)= .972; $F(2,344)=4.95$, $p < .05$). Bu ortak etkiye göre problem çözme puanlarında anlamlı bir farklılık görülürken ($F(1,345) = 7.002$; $p < .05$), eleştirel düşünme puanlarında anlamlı farklılık görülmemiştir. Ortalamalardan sosyal bilimlerde okuyan bayanların fen bilimlerinde okuyan bayanlardan daha yüksek bir problem çözme becerisine sahip oldukları anlaşılmıştır. Buna karşın erkeklerde tam tersi bir durum olduğu görülmüştür. Bir başka deyişle fen bilimlerinde okuyan erkeklerin sosyal bilimlerde okuyan erkeklerden daha yüksek bir problem çözme becerisine sahip oldukları saptanmıştır. Cinsiyet, sınıf düzeyi ve eğitim alanı ortak etkileşiminin problem çözme ve eleştirel düşünme puanları üzerinde anlamlı bir farklılık yaratmadığı anlaşılmıştır.

Araştırmanın Sonuçları: Yapılan korelasyon analizi öğrencilerin eleştirel düşünme eğilimleri ile problem çözme becerileri arasında anlamlı bir ilişki olduğunu göstermiştir. Bu çalışmada cinsiyetin öğrencilerin eleştirel düşünme eğilimleri ve problem çözme becerileri açısından anlamlı bir fark yaratmadığı anlaşılmıştır. Genel olarak literatür tarandığında bu çalışmadan elde edilen bulguları destekleyen sonuçlara ulaşılmıştır. Üniversite öğrencilerinde eleştirel düşünme becerilerinin cinsiyet açısından farklılaşıp farklılaşmadığının incelendiği birçok çalışmada bu iki değişken açısından anlamlı bir farklılaşmanın olmadığı belirlenmiştir. Aynı

ekilde aratırmalar ğrencilerin problem özme becerileri ile cinsiyetleri arasında da anlamlı bir ilikinin olmadığını göstermiştir.

Ayrıca, bu aratırmada birinci sınıf ğrencileri ile drdnc sınıf ğrencileri arasında eletirel dnme eęilimleri ve problem özme becerileri aısından da anlamlı bir farklılık ortaya çıkmıştır. Birinci sınıfta okuyan ğrenciler gerek eletirel dnme, gerekse problem özme becerisi yönünden kendilerini, drdnc sınıfta okuyan ğrencilere göre daha yetersiz olarak algıladıklarını bildirmişlerdir. Durumun son sınıf ğrencilerinin lehine olması, verilen drt yıllık üniversite eęitiminin eletirel dnme ve problem özme becerisi üzerinde olumlu bir etkisi olduęu eklinde yorumlanabilir. ğrencilerin eęitim gördükleri alanlara göre eletirel dnme puanları anlamlı bir ekilde farklılaırken problem özme puanlarında ise anlamlı bir farklılığın olmadığı saptanmıştır. Cinsiyet ve sınıf düzeyinin ortak etkileimi ile sınıf düzeyi ve eęitim alanının ortak etkileimini her iki baęımlı deęiken üzerinde anlamlı bir fark yaratmamıştır. Cinsiyet ve eęitim alanındaki ortak etkiye göre problem özme puanlarında anlamlı bir farklılık görülürken, eletirel dnme puanlarında anlamlı farklılık görülmemiştir. Cinsiyet, sınıf düzeyi ve eęitim alanı ortak etkileiminin problem özme ve eletirel dnme puanları üzerinde anlamlı bir farklılık yaratmadığı anlaılmıştır.

neriler: niversitelerdeki psikolojik danıma ve rehberlik servislerinin ğrencilerin eletirel dnme eęilimleri ve problem özme becerilerini gelitirmelerine yönelik grup rehberlięi etkinliklerini arttırmaları, ğretim elemanlarının derslerinde bu becerileri kazanmalarına olanak saęlayacak ders etkinliklerine yer vermeleri yararlı olabilir

Anahtar Szckler: California eletirel dnme eęilimleri envanteri, eletirel dnme, problem özme envanteri, üniversite ğrencileri

Investigation of University Students' Depression*

A.Aykut Ceyhan**
Esra Ceyhan***
Yıldız Kurtyılmaz****

Suggested Citation:

Ceyhan, A.A., Ceyhan, E., & Kurtyılmaz, Y. (2009). Investigation of university students' depression. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 75-90.

Abstract

Problem Statement: In recent years, the number of university students diagnosed, treated, and given medication for depression has increased. It can be stated that depression is a serious emotional problem for university students because it affects younger and older adults directly and profoundly. It leads to loss of interest and pleasure in ordinary activities, great sadness, feelings of guilt and worthlessness, loss of appetite and sleep, etc. Thus depression decreases the quality of an individual's life.

Purpose of Study: This study was carried out to investigate the depression levels of university students according to their gender, body image satisfaction, residence location types, and academic achievement.

Methods: Participants consisted of 440 university students attending Anadolu University, Turkey. Data was collected through the Beck Depression Inventory, the Pleasure of Body Organs Scale, and the Personal Information Questionnaire. For data analysis, descriptive statistics, the independent-samples t test, the two-way ANOVA, and the two-way ANCOVA were used.

Findings and Results: It was found that university students who were dissatisfied with their body image were more depressed than those who were satisfied with their body image. Moreover, university students who had a lower academic achievement level demonstrated significantly more depressive symptoms than those who demonstrated high academic achievement. It was indicated that the depression levels of university students did not significantly differ according to gender and residence location types and their interaction, and the same findings for the two factors were obtained when body image satisfaction and academic achievement levels were held constant.

* This manuscript was presented at Paris International Conference on Education, Economy & Society, Paris, 17-19 July 2008

** Corresponding Author, Assoc.Prof.Dr., Anadolu University, Faculty of Education, aceyhan@anadolu.edu.tr

*** Assos.Prof.Dr., Anadolu University, Faculty of Education, ecyhan@anadolu.edu.tr

**** M.A., Anadolu University, Faculty of Education, ykurtyilmaz@anadolu.edu.tr

Conclusions and Recommendations: Since depression is a major psychological problem for university students, it can be stated that effective psychological help services should be rendered to university students. Thus, to improve coping skills against depression, various programs should be built for increasing their self-acceptance levels, enhancing their communication skills, and improving social relations. In these programs, factors of dissatisfaction with physical appearance and academic issues should be taken into consideration. This study has had some limitations: because of the limited number of participants, the interactions among depression levels and gender, body image satisfaction and residence location types should be examined in larger groups. Qualitative studies can be carried out to get a more detailed picture of university students' depression.

Keywords: Academic achievement, residence location type, body image satisfaction, depression, gender, university student

During the years at a university, students can become perplexed because they are in a period of independent identity exploration, instability, possibilities, feeling in-between, and they are self-focused. They start to explore various roles, lifestyles, and worldviews. University students try to establish a balance between meeting their own expectations and meeting the expectations of others, such as friends and family. They feel that they are in between adolescence and adulthood. At the same time, they explore themselves and self exploration leads to a rise and fall in self esteem (Arnett, 2004). Therefore, university life brings about both the possibilities and the negative and challenging life conditions. As a result, university students may become more vulnerable to some psychological symptoms such as anger, anxiety, and depression. Therefore, the prevalence of some psychopathologies in university students is high (Rosenthal & Schreiner, 2000). Depression is an especially frequently experienced problem among university students. This is confirmed by a variety of studies which indicated that the rate of depression among university students was 21% (Ceyhan, Ceyhan, & Kurtyılmaz, 2005), 25.4% (Aylaz, Kaya, Dere, Karaca, & Bal, 2007), 26.2% (Özdel, Bostancı, Özdel, & Oğuzhanoglu, 2002), 35% (Gökçakan & Gökçakan, 1998), and 35.2% (Çam Çelikel & Erkorkmaz, 2008).

Depression is an emotional state revealing itself with loss of interest and pleasure in ordinary activities, great sadness, feelings of guilt and worthlessness, loss of appetite and sleep, and loss of sexual desire (Davison & Neale, 1994). Therefore, depression decreases the quality of an individual's life (Demyttenaere, De Fruyt & Huygens, 2002). For this reason, there are lots of studies trying to understand the nature of depression and the relationship between depression and the variables thought to be related with depression (Aylaz et al, 2007; Çam Çelikel & Erkorkmaz, 2008; Dorahy et al, 2000).

In the literature, it was observed that individuals' depression levels have been related with their body satisfaction (i.e. satisfaction with one's physical appearance). The findings revealed that body dissatisfaction was related with depression not only in clinical female samples (Wiederman & Pryor, 2000) but also in nonclinical female

samples (Denniston, Roth & Gilroy, 1992). Moreover, depression was found as a significant predictor of body dissatisfaction (Joiner, Wonderlich, Metalsky, & Schmidt, 1995).

Extreme concerns about body image may lead individuals to become overly involved in their body shape and may increase the probability of being dissatisfied with their own body image. Therefore, in the studies investigating the relationship between depression and body image (dis)satisfaction, the variable or weight preoccupation, or a drive for thinness, was also taken in to account. In one of these studies, it was found that weight preoccupation resulted in depression (Ledoux, Choquet & Flament, 1991). Also, drive for thinness was found to be related with depression (Willcox & Sattler, 1996). In another study (Cohen-Tovée, 1993), it was found that depression caused individuals to be more severely overconcerned with their body shape and weight. In contrast, the participants who were not overconcerned with their body shape and weight were not affected by induced depression. In sum, it can be stated that a depressed mood lead females to become overconcerned with their bodies and causes them to be dissatisfied with their bodies (Cohen-Tovée, 1993). Parallel to this finding, body image dissatisfaction in males was also found to be related with depression (Olivardia, Pope, Borowiecki & Cohane, 2004). Similarly, sociotropy, body dissatisfaction, and drive for thinness were found to predict depression significantly and the model did not change according to gender, both females and males strive to reach an ideal body (Oates-Johnson & DeCourville, 1999).

Academic achievement is another variable taken in to account in depression studies. There was a significant relationship between academic impairment and depression in university students. Academic impairment was more pervasive among depressed university students, and as depression becomes more severe, academic impairment rises (Heiligenstein & Guenther, 1996). In a qualitative study, participants stated that their academic problems were among the causes of their depression (Anderson, 2003). This qualitative finding was supported by some research findings indicating that depression level and academic achievement of university students were negatively related (Aydın, 1988; Çam Çelikel & Erkorkmaz, 2008; Koç, Avşaroğlu & Sezer, 2004; Özdel et al, 2002).

One of the most studied variables for depression is gender, and the findings have been inconsistent and contradictory (Rosenthal & Schreiner, 2000). In some of these studies, females were found to be more depressed than males (W.E. Kelly, K.E. Kelly, Brown & H.B. Kelly, 1999; McCarty et al, 2008; Oates-Johnson & DeCourville, 1999; Rosenthal & Schreiner, 2000; Wiseman, Gutfreund & Lurie, 1995). In contrast, some studies have obtained no difference in terms of gender (Aydın, 1988; Aydın & Demir, 1989; Aylaz et al, 2007; Çam Çelikel & Erkorkmaz, 2008; Dorahy et al, 2000; Gökçakan & Gökçakan, 1998; Heiligenstein & Guenther, 1996; Hisli, 1988; Özdel et al, 2002).

In addition to the variables of body image satisfaction, academic achievement, and gender, studies have investigated the residence location types of the university students in connection with depression. Some of these studies have supported the

statement that there is a relationship between residence location and depression of university students. It was found that university students staying with their families have fewer symptoms of depression than the students staying away from their families (Aydın & Demir, 1989). In a qualitative study, participants explained that living in a dormitory was one of the causes of their depression; some characteristics of the dormitory make them more vulnerable to depression (Anderson, 2003). For instance, they found the dormitory building was dreary and more crowded. Most of the students complained about the monoculture in the dormitory. They become bored because they are always together with peers having similar problems, similar expectations, similar anxieties, etc. Some of the participants stated that they do not come across any adult figure at dormitories. These characteristics of dormitories were reposted by college students as depressing (Anderson, 2003). On the other hand, there are some findings indicating that there is no relationship between depression and residence location (Özdel et al, 2002; Patchett, 2005).

Depression is the most frequent psychological distress that university students face. Because of the prevalence of depression among university students, its investigation is of prominence (Aylaz et al, 2007; Oates-Johnson & DeCourville, 1999; Özdel et al, 2002). For this reason, this study aimed to investigate the depression levels of university students in terms of gender, body image satisfaction, academic achievement, and residence location type. Thus, answers to the following research questions were sought.

- (1) What is the depression level of the university students?
- (2) Do depression levels of university students significantly differ according to their body image satisfaction levels?
- (3) Do depression levels of university students significantly differ according to their academic achievements?
- (4) Do depression levels of university students significantly differ according to gender and residence location types?
- (5) Do depression levels of university students significantly differ according to gender and residence location types, holding constant the body image satisfaction and academic achievement?

Method

Participants

The study was conducted with 440 volunteer university students attending various programs (English Language Teaching, French Language Teaching, Primary School Mathematics Teaching, Primary School Education, Arts and Crafts Education, Preschool Education, Computer Education and Instructional Technology, and Education of the Mentally Disabled) at Anadolu University, Turkey. Females participants constituted 71.60% (315), and males constituted 28% (123). 102 were

freshman (23.20%), 105 were sophomores (23.90%), 110 were juniors (25%), and 123 were seniors (28%). The mean of their cumulative GPAs was 2.92 on a 4 point scale.

Instruments

Beck Depression Inventory (BDI). The BDI was developed by Beck, Ward, Mendelson, Mock, and Erbaugh in 1961 for measuring emotional, somatic, cognitive and motivational symptoms of depression. This scale is not used for making a diagnosis of depression but for determining the levels of depression symptoms. The scale consists of 21 items related to symptom categories of depression. The higher the points obtained, the higher the level and severity of the depression (Savaşır & Şahin, 1997).

A number of reliability and validity studies were carried out by various researchers for the adaptation of the BDI-78 for the Turkish population. In studies conducted by Hisli (1988, 1989) the correlation coefficient between the Turkish and English form of the BDI was found as .81 and .73, the split-half reliability was .74, and the correlation coefficients between the BDI and the MMPI-D (Minnesota Multiphasic Personality Inventory-Depression) in a psychiatric patient sample and a university students sample were 0.63 and 0.50, respectively. The correlation coefficient between the BDI and the STAI-T (State-Trait Anxiety Inventory-Trait Subscale) was obtained as .55. Şahin, Rugancı, Taş, Kuyucu, and Sezgin (1992) computed the correlation coefficient between the BDI and Ineffective Coping as .50, and found the correlation coefficient between the BDI and Effective Coping as -.30. Moreover, Hisli (1988) indicated that the point obtained from the BDI is 17 and over differentiates the level of depression requiring treatment with over 90% confidence.

In most studies in which the BDI was used (Aylaz et al, 2007; Ceyhan et al, 2005; Çam et al, 2008; Gökçakan & Gökçakan, 1998; Özdel et al, 2002) it has been seen that a score of 17 points and above was accepted as the cut off point. Moreover, the cut off point regarding the severity of depression was also designated by Bumberry, Oliver & McClure (1978, as cited in Wiseman et al, 1995). The points between 0 and 9 were classified as 'non-depressed', the points between 10 and 15 were classified as 'mildly depressed', the points between 16 and 23 were classified as 'moderately depressed', and 24 points or more was classified as 'severely depressed'.

Pleasure of the Body Organs Scale. The Pleasure of the Body Organ scale was developed by Çetinkaya (2004) and assesses ones satisfaction with their body organs. The scale is a-7 point Likert type scale and it consists of 17 items. The points system of the scale ranges from 17-119. The higher the score received, the higher the satisfaction with the body. In the validity study, the correlation coefficient between the Pleasure of the Body Organs Scale and the Body Image Quality of Life Inventory was found as $r = .33, p < .01$. In the study, the scale was applied to 531 participants and the cronbach alpha was found as .83. In the validity studies, the correlation coefficient of the scale with the Body Image Quality of Life Inventory was found as $r = .40, p < .01$ (Çetinkaya, 2004).

Personal Information Questionnaire. The questionnaire used was developed specifically for this study. The questionnaire included various questions about university students'

gender (female/male), grade (freshman/sophomore/junior/senior), cumulative Grade Point Average (CumGPAs), and residence location type (staying with family/staying in dormitory/staying with friends at home/staying alone at home/other).

Data Collection and Analysis

Data was collected in April and May 2008 during classes and analyzed by means of the SPSS. For data analysis, descriptive statistics, independent-samples t test, the two-way ANOVA, and the two-way ANCOVA were used. The homogeneity of variance was controlled with the Levene test. A significance level of .05 was accepted for all analyses.

In this research, to examine the relationship between depression and body image satisfaction and between depression and academic achievement, it was necessary to first determine the extreme groups for the variables of body image satisfaction and academic achievement. Subjects with the scores at either extreme ends were selected. The selection criterion was a +/- 1 standard deviation from the mean. Thus, subjects with relatively higher scores formed the "high" group; subjects with relatively lower scores formed the "low" group. The means and standard deviation of the points gathered from the pleasure of the body organs scale were first calculated in order to divide the participants into two groups: "body image satisfaction" and "body image dissatisfaction". According to this calculation, the mean and standard deviation of scale scores were 90.55 and 15.26, respectively. The sample was divided into a "body image satisfaction group" which consisted of 80 subjects whose points were 106 and over, and into a "body image dissatisfaction group" comprising 77 subjects whose points were 75 and lower. 283 subjects whose scores lay between 75 and 106 were not included in the analysis. Rather, data obtained from 157 students for the variable of body image satisfaction were analyzed.

To divide the participants into two groups as high and low academic achievers, a means and standard deviation of cumulative GPAs were calculated. The mean and standard deviation of cumulative GPAs were 2.93 and 0.40, respectively. According to these values, one standard deviation below the mean was 2.53 and one standard deviation above was 3.33. Participants were grouped in a "low academic achievement group" which consisted of 74 participants whose cumulative GPAs were between 1.66 and 2.53, and a "high academic achievement group" consisting of 78 participants whose scores fell between 3.33-3.80. 288 subjects whose score were between 2.53 and 3.33 were not included in the analysis. Therefore, the data obtained from 152 students was analyzed.

Findings and Results

1. The findings concerning depression levels of university students

Obtained descriptive statistics indicated that the mean of the university students' depression levels was 12.63 and the standard deviation of the distribution was 8.21. When the distribution of depression points were examined, 25% of the university students at 7.00 or below, 50% of the participants' points were between 8.00 and 11.00, and 75% of the participants' points were between 12.00 and 17.00. According to

the cut-off point of 17.00 obtained from the BDI (Hisli, 1989), 26.81% (118) of the current research participants have severe depression symptoms. Furthermore, if the distribution of the participants was examined in terms of other suggested cut off points regarding the severity of depression (Bumberry et al, 1978 as cited in Wiseman et al, 1995), it can be stated that 177 (40.22%) participants had no depression, 130 participants (29.55%) were mildly depressed, 90 of them (20.46%) were moderately depressed, and 43 of them (9.77%) were severely depressed.

2. The findings concerning body image satisfaction

According to the criterion explained in the method section, the participants were grouped into a "body image satisfaction group" and a "body image dissatisfaction group." The depression levels of these two groups were compared by an independent-samples t test. The findings indicated that the university students who were dissatisfied with their body image (N=77, M = 17.91, SD = 9.18) were significantly more depressed than those who were satisfied with their body image (N=80, M = 8.38, SD = 6.95), $t(155) = 7.34, p = .001$ (two-tailed), $\eta^2 = .26$.

3. The findings concerning academic achievement

Whether depression levels of university students differ according to academic achievement levels was also investigated. According to the criterion explained in the method section, the participants were grouped into a "low academic achievement group" and a "high academic achievement group." The analysis carried out with the independent-samples t test indicated that the university students who had low academic achievement (N=74, M = 14.59, SD = 9.90) demonstrated significantly more depressive symptoms than those who had high academic achievement (N=78, M = 11.10, SD = 7.25), $t(150) = 2.57, p = .01$ (two-tailed), $\eta^2 = .04$.

4. Findings concerning gender and residence location types

To determine whether depression levels of females and males change as a function of residence location type, the mean points and standard deviations related with these variables are found as shown in Table 1.

Means and Standard Deviations of Depression Levels for Female and Male University Students

Residence location Types	Mean		Standard deviation		Size	
	Female	Male	Female	Male	Female	Male
Staying with family	11.93	15.32	6.90	11.53	59	19
Staying in dormitory	12.81	11.39	8.42	8.52	147	28
Staying with friends at home	11.57	13.43	6.46	8.22	73	72

Table 1: "other" (n=29), "stay alone at home"(n=11) and missing data (n=2) are omitted

To determine whether these observed differences were statistically significant, a two way analysis of variance (ANOVA) procedure was carried out. The assumptions underlying a two-way ANOVA analysis were initially investigated. Firstly, Levene's test

verified the assumption of equality of variance, $F(5,392) = 1.71, p = .132$. In the framework of this finding, a two-way ANOVA was conducted. The results indicated that the depression levels of university students did not differ significantly according to the main effects of gender, $F(1,392) = 1.62, p = .20$, residence location types, $F(2,392) = .66, p = .52$, or according to their interaction, $F(2,392) = 1.89, p = .15$.

5. Findings concerning gender and residence location types holding body image satisfaction and academic achievement constant

In addition, to determine whether depression levels of university students significantly differ according to gender and residence location types, holding body image satisfaction level and academic achievement level (or cumGPAs) constant, a two way analysis of covariance (ANCOVA) was performed. For this purpose, the assumptions underlying a two-way ANCOVA analysis were initially investigated. Generally, the depression variable is distributed throughout the sample for any specific value of the covariates (body image satisfaction scores and academic achievement level) for each level of gender and residence location types when examined by plots.

Since two covariates were used, the correlation with one another was checked and found as insignificant ($r = -.06$). It was observed that there was a linear relationship between the depression level and the two covariates of gender and residence location types when scatterplots were examined. In addition, the homogeneity of regression slopes, including the relationship between the covariates, and the depression variable for each of gender and residence location types groups was also investigated by scatterplots graphically and through a univariate analysis procedure. Thus, before conducting a two-way ANCOVA, the homogeneity-of-slopes assumption was tested. In the output obtained from the analysis, the significance level for the interactions was found not to be statistically significant [$F(1,371) = 1.10, p = .29$ for gender and academic achievement level; $F(1,371) = .99, p = .32$ for gender and body image satisfaction; $F(2,371) = .21, p = .81$ for residence location type and academic achievement level; $F(2,371) = .45, p = .64$ for residence location type and body image satisfaction; $F(1,371) = 1.18, p = .28$ for gender, academic achievement level, and body image satisfaction; and $F(1,371) = .28, p = .75$ for residence location type, academic achievement level, and body image satisfaction]. Moreover, Levene's test verified the assumption of an equality of variance [$F(5,381) = 2.18, p = .055$]. Based on these results, the two-way ANCOVA analysis was carried out to explore the main effects of gender and residence location types and the interaction effects.

The results of the ANCOVA analysis showed that the depression levels of university students did not significantly differ according to the main effects of gender and residence location types and their interaction. As a result, the students' depression levels did not vary according to gender and residence location types, or the interaction between them [$F(1,379) = 3.55, p = .06$ for main effect of gender, $F(2,379) = .32, p = .73$ for main effect of residence location types, and $F(2,379) = .008, p = .99$ for their interaction].

Discussion

According to the findings, the mean level of university students' depression was 12.63 and the standard deviation of the distribution was 8.21. Similarly, the other studies have also found the mean and standard deviation of university students' depression as 12.72 and 8.26 (Ceyhan et al, 2005), 12.8 and 7.19 (Özdel et al, 2002), 13.47 and 8.45 (Aylaz et al, 2007), 15.54 and 8.15 (Çam et al, 2008), respectively. When these values are taken in to account, it can be stated that the mean of the current study is quite similar to the means of previous studies.

According to the BDI cut off points for the depression severity (Bumberry, Oliver & McClure, 1978, as cited in Wiseman et al, 1995), 40.22% of the participants have no depression. 29.55% of them are mildly depressed, 20.46% of them are moderately depressed, and the remaining 9.77% of them are severely depressed. When the distribution of the current study is examined in terms of a 'point of 17 or above' criterion (Hisli,1989), 26.81% of the participants have severe depression. This rate is also similar with other research rates changing between 21% and 35.2% (Aylaz et al, 2007; Ceyhan et al, 2005; Çam et al, 2008; Gökçakan & Gökçakan, 1998; Özdel et al, 2002). Thus, these rates indicated that, for most of the university students, depression is a significant problem that requires getting psychological help.

Although the cut off points for the severity and level of depression were accepted as the criteria, these values should be taken cautiously. It is based on the individuals expression about their depressive symptomatology during the previous week. Therefore, depressive symptomatology may be transient. Despite this, points from the BDI give information about proneness to depression. Taking this in to account, the need for more psychological help to university students becomes apparent.

The current findings also denoted that university students dissatisfied with their body image have higher depression levels than those satisfied with their body image. This finding is consistent with the findings in other literature (Cohen-Tovée, 1993; Denniston et al, 1992; Joiner et al, 1995; Ledoux et al, 1991; Oates-Johnson & DeCourville, 1999; Olivardia et al, 2004; Wiederman & Pryor, 2000; Willcox & Sattler, 1996, Yanarca Uras, 2004). The relationship between body dissatisfaction and depression can be explained by low self esteem (de Man, Gutierrez, & Sterk, 2001) and the overconcern with body shape (Cooper & Taylor, 1988, as cited in Wiederman & Pryor, 2000) or weight preoccupation (Cohen-Tovée, 1993; Oates-Johnson & DeCourville, 1999). Weight preoccupation causes individuals to be hypercritical about their own body and physical appearance. Weight preoccupation may be more likely to affect some individuals who have a low self-esteem and are sociotropic, because they tend to have a need for approval and acceptance from others, in order to accept themselves. Therefore, it may cause individuals become more vulnerable to depression (Oates-Johnson & DeCourville, 1999).

Individuals who are sociotropic and have a low self esteem feel themselves more valuable when they are approved of and loved by others. Thus, in order to feel valuable and to be approved in terms of physical appearance, they compare their

bodies with the accepted ideal body form of the society. When they think that they can not reach the ideal body form, they may have a feeling of failure and their self esteem may shrink. In turn, for the evaluations used in these studies, they make a negative self evaluation which is a trivet of the negative triad used by Beck to explain depression (de Man et al, 2001). These negative feelings of the individuals, who evaluate themselves in terms of physical appearance, may be generalized and decrease self-esteem and result in depression. This would help explain why university students who are dissatisfied with their body images have higher depression levels than those who are relatively more satisfied with their body images.

The other finding of this research is that depression levels of the individuals with low academic achievement are higher than that of those with high academic achievement. This research finding is consistent with other studies (Aydın, 1988; Çam Çelikel & Erkorkmaz, 2008; Koç, Avşaroğlu & Sezer, 2004; Özdel et al, 2002; Yeh et al, 2007). Since university students need positive reinforcement to be motivated, considering the period of life that they are in, they avoid negative situations (Gemelli, 1996 as cited in Yeh et al, 2007) and academic problems constitute one of these negative experiences.

There is a negative relationship between academic achievement and depression. This is because, according to the life stress model, the desire for academic achievement leads individuals to feel under intensive stress (Beck et al, 1983, as cited in McCarty et al, 2008). Both the individual and his/her family expect that the individual be successful (Anderson, 2003). The student who can not meet the expectation of achievement and can not achieve this goal is more likely to sink into depression (Patchett, 2005). Thus, academic underachievement is accepted as one of the causes of depression. On the other hand, academic achievement facilitates adaptation to university life (Baker & Siryk, 1984). For this reason, academically low achievers may have a higher probability to experience adaptation difficulties and depression. Even after past studies have been investigated, it is difficult to state whether academic achievement is an antecedent or consequence of depression. The same is valid for this study. Therefore, it can be stated that these two variables are in a reciprocal relationship.

In this study, depression levels of university students did not significantly differ according to the main effects of gender and residence location types and their interaction. The finding that there is no difference between genders in terms of depression is consistent with the findings of other research (Aydın, 1988; Aydın & Demir, 1989; Aylaz et al, 2007; Ceyhan et al, 2005; Çam Çelikel & Erkorkmaz, 2008; Dorahy et al, 2000; Gökçakan & Gökçakan, 1998; Heiligenstein & Guenther, 1996; Hisli, 1988; Özdel et al, 2002). On the other hand, it is inconsistent with the findings which indicated that females were more depressed than males (Kelly et al, 1999; McCarty et al, 2008; Oates-Johnson & DeCourville, 1999; Rosenthal & Schreiner, 2000; Wiseman et al, 1995). All these results indicate that there is no clear agreement in research so far comparing that depression levels of females and males (Rosenthal & Schreiner, 2000). Finding no difference between genders may be caused by the nature

of the period in which university students are. Gender differences may diminish as a result of education. Moreover, university students are open minded about gender appropriate behaviors. Differences between females and males caused by gender role stereotypes have lessened. At the same time, problems for females and males attending university are more common and similar. Hence, females and males did not differ in terms of depression levels.

Depression levels of university students also did not differ according to residence location type, and at the same time no significant difference was found between female and male depression levels according to their residence location. This finding is inconsistent with some research findings (Aydın & Demir, 1989), but finding no effect was consistent with most studies (Özdel et al, 2002; Patchett, 2005). In this study, the Negative Mood Regulation (NMR) and overall adjustment were found to be significant predictors of depression. Students with high NMR easily cope with the changes in living environment and expectations caused by transition into a university setting regardless of residence (Patchett, 2005). of the lack of residence having an effect may be caused by the fact that participants can control their negative feelings induced by their residence type location. Participants being able adapt to their living environment easily regardless of residence location types, may be why no difference was found in terms of residence location type.

Moreover, university students listed loneliness and lack of social relationships among the main causes of their depression (Anderson, 2003). This view was also supported by the finding that loneliness is the main predictor of depression (Patchett, 2005). For these reasons, university students need to get social support in order to cope with loneliness, to protect themselves against depression, and to adapt to their environment. While the participants living with friends and in dormitories may get social support from their friends, the participants at home might get social support from their families. Therefore, since all the participants may not feel lonely, no difference was observed among the participants from different residence location types.

Furthermore, all the university students pass through transition periods although their experiences are very different. Despite the fact that the fields that they have to adapt to vary, t all university students experience changes and they have to adapt to these changes. For example, university students staying with their families want to establish intimate relationships with their peers but they are also obliged to inform their families about their plans (when they will be at home, etc.). On the other hand, even though the university students staying away from their families experience more freedom, they cope with various problems such as food issues, rejection from roommates, getting along with roommates, etc.

All university students' relationships with their families change wherever they reside. Moreover, the influence of families on young adults may diminish, so a difference might not be observed on depression levels in terms of residence type. As a result of this research, it was revealed that depression levels of university students did not differ according to gender, residence location type, or the interaction between

these two variables, even when the variables of body image satisfaction and academic achievement were controlled. This finding indicated that gender and residence are not important factors. Conversely, satisfaction with physical appearance is an important factor making a difference. University students who are dissatisfied with their physical appearance have been found to have higher depression levels. Furthermore, depression levels of low academic achievers were found to be significantly higher than that of high academic achievers. It can be stated that body image satisfaction and academic achievement are the most significant factors contributing to depression in university students.

Since depression is a major psychological problem for university students, university should establish effective psychological help services for the students. Various programs for the prevention of depression and coping with depression can be built. In these programs students' self-acceptance levels can be heightened, their communication skills can be enhanced, and social relations can be improved. In the designing of these programs, factors of dissatisfaction with physical appearance and academic issues must be taken into consideration. Moreover, seminars, group guidance and counseling about time management, effective studying techniques, etc. can be carried out.

The interactions among depression levels and gender, body image satisfaction and residence location type needs to be further examined to include a larger selection of participants. Perhaps one change could also be to ask participants about the extent to which they are pleased with their place of residence. Various factors such as self-esteem, loneliness, and sociotropy may be added to the research design. Qualitative studies can also be carried out to get a more detailed picture of university students' depression.

References

- Anderson, M. (2003). The experience of depression in first-year students living in residence: A phenomenological inquiry. Unpublished doctoral dissertation, University of Victoria, Canada.
- Arnett, J. J. (2004). *Emerging adulthood: The winding road from the late teens through the twenties*. New York: Oxford University Press.
- Aydın, G. (1988). Üniversite öğrencilerinde depresyon, açıklama biçimi ve akademik başarı ilişkisi. [The relationships among university students' depression, depressive explanatory style and academic achievement]. *Psikoloji Dergisi*, 6 (22), 6-13.
- Aydın, G. & Demir, A. (1989). O.D.T.Ü. öğrencilerinde depresif belirtilerin yaygınlığı. [Prevalance of depressive symptoms among METU students]. *İnsan Bilimleri Dergisi*, 8(1), 27-40.
- Baker, R. W. & Siryk, B. (1984). Measuring adjustment to college. *Journal of Counseling Psychology*, 31(2), 179-189.

- Ceyhan, E., Ceyhan, A. A., & Kurtyılmaz, Y. (2005). Depression among Turkish female and male university students. *Social Behavior and Personality*, 33(4), 329-340.
- Cohen- Tovée, E. M. (1993). Depressed mood and concern with weight and shape in normal young women. *International Journal of Eating Disorders*, 14(2), 223-227.
- Çam Çelikel, F. & Erkorkmaz, Ü. (2008). Üniversite öğrencilerinde depresif belirtiler ve umutsuzluk düzeyleri ile ilişkili etmenler [Factors related to depressive symptoms and hopelessness among university students]. *Nöropsikiyatri Arşivi*, 45, 122-129.
- Çetinkaya, H. (2004). Beden imgesi, beden organlarından memnuniyet, benlik saygısı, yaşam doyumunu ve sosyal karşılaştırma düzeylerinin demografik değişkenlere göre farklılaşması [Differentiation of body image, pleasure of body organs, self-esteem, life satisfaction and social comparison level according to demographic variables]. Yayınlanmamış Yüksek Lisans Tezi, Mersin Üniversitesi.
- Davison, G. C., & Neale, J. M. (1994). *Abnormal Psychology*. New York: John Wiley & Sons, Inc.
- de Man, A. F., Becerril Gutierrez, B. I., & Sterk, N. (2001). Stability of self-esteem as moderator of the relationship between level of self-esteem and depression. *North American Journal of Psychology*, 3(2), 303-308.
- Demyttenaere, K., De Fruyt, J., & Huygens, R. (2002). Measuring quality of life in depression. *Current Opinion in Psychiatry*, 15(1), 89-92.
- Denniston, C., Roth, D., & Gilroy, F. (1992). Dysphoria and body image among college women. *International Journal of Eating Disorders*, 12 (4), 449-452.
- Dorahy, M. J., Lewis, C. A., Schumaker, J. F., Akuamoah-Boateng, R., Duze, M.C., & Sibiya, T. (2000). Depression and life satisfaction among Australian, Ghanaian, Nigerian, Northern Irish, and Swazi University Students. *Journal of Social Behavior & Personality*, 15 (4), 569-580.
- Gökçakan, Z., & Gökçakan, N. (1998). Öğretmen adaylarında depresyon. [Depression in teacher trainees]. 7. Ulusal Eğitim Bilimleri Kongresi Bildirisi, Konya
- Heiligenstein, E. & Guenther, G. (1996). Depression and academic impairment in college students. *Journal of American College Health*, 45 (2), 59-64.
- Hisli, N. (1988). Beck Depresyon Envanteri'nin geçerliği üzerine bir araştırma [A research about the validity of Beck Depression Inventory]. *Psikoloji Dergisi*, 6(22), 118-126.
- Hisli, N. (1989). Beck Depresyon Envanteri'nin üniversite öğrencileri için geçerliği, güvenilirliği [Reliability and Validity of Beck Depression Inventory for university students]. *Psikoloji Dergisi*, 7 (23), 3-13.
- Joiner, T. E., Wonderlich, S. A., Metalsky, G. I., & Schmidt, N. B. (1995). Body dissatisfaction: A feature of bulimia, depression, or both? *Journal of Social and Clinical Psychology*, 14 (4), 339-355.
- Kelly, W. E., Kelly, K. E., Brown, F. C., & Kelly, H. B. (1999). Gender differences in depression among college students: A multicultural perspective. *College Student Journal*, 33 (1), 72-77.
- Koç, M., Avşaroğlu, S., & Sezer, A. (2004). Üniversite öğrencilerinin akademik başarıları ile problem alanları arasındaki ilişki [Relationship between academic success and problematic areas of university students]. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 11, 483-498.

- Ledoux, S., Choquet, M., & Flament, M. (1991). Eating disorders among adolescents in an unselected French population. *International Journal of Eating Disorders*, 10(1), 81-89.
- McCarty, C. A., Mason, W. A., Kosterman R., Hawkins, J. D., Lengua, L. J., & McCauley, E. (2008). Adolescent school failure predicts later depression among girls. *Journal of Adolescent Health*, 43 (2), 180-187.
- Oates-Johnson, T. & DeCourville, N. (1999). Weight preoccupation, personality, and depression in university students: An interactionist perspective. *Journal of Clinical Psychology*, 55 (9), 1157-1166.
- Olivardia, R., Pope, H. G., Borowiecki, J. J., & Cohane, G. H. (2004). Biceps and body image: The relationship between muscularity and self-esteem, depression, and eating disorders symptoms. *Psychology of Men and Masculinity*, 5 (2), 112-120.
- Özdel, L., Bostancı, M., Özdel, O., & Oğuzhanoğlu, N.K. (2002). Üniversite öğrencilerinde depresif belirtiler ve sosyodemografik özelliklerle ilişkisi [The relationship with sociodemographic characteristics and depressive symptoms in university students]. *Anadolu Psikiyatri Dergisi*, 3, 155-161.
- Patchett, E. (2005). Negative mood regulation expectancies and residence location as predictors of college students' adaptation to college, depression and loneliness. Unpublished doctoral dissertation, California State University.
- Rosenthal, B. S., & Schreiner, A. C. (2000). Prevalence of psychological symptoms among undergraduate students in an ethnically diverse urban public college. *Journal of American College Health*, 49, 12-18.
- Savaşır I., & Şahin, N. H. (1997). Bilişsel-davranışçı terapide değerlendirme: Sık kullanılan ölçekler [The evaluation in the cognitive-behavioral therapy: The scales used frequently]. Ankara: Türk Psikologlar Yayını.
- Şahin, N. H., Rugancı, N., Taş, Y., Kuyucu, S., & Sezgin, N.(1992). Üniversite öğrencilerinde stresle ilgili faktörler ve etkili başa çıkma [Stres-related factors and effective coping among university students]. 25. Psikoloji Kongresi Bildirisi, Brüksel.
- Wiederman, M. W. & Pryor, T. L. (2000). Body dissatisfaction, bulimia, and depression among women: The mediating role of drive for thinness. *International Journal of Eating Disorders*, 27(1), 90-95.
- Willcox, M., & Sattler, D. N. (1996). The relationship between eating disorders and depression. *The Journal of Social Psychology*, 136 (2), 269-271.
- Wiseman, H., Gutfreund, D.G., & Lurie, I. (1995). Gender differences in loneliness and depression of university students seeking counseling. *British Journal of Guidance & Counseling*, 23 (2), 231-245.
- Yanarca Uras, B. (2004). Body satisfaction and depression symptoms relationship among university students: The mediating role of self esteem [Beden imgesi ve depresyon semptomları ilişkisi: Özgüvenin, bu ilişkide oynadığı aracı rol]. Yayınlanmamış Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi.
- Yeh, Y. C., Yen, C. F., Lai, C. S., Huang, C. H., Liu, K. M. & Huang, I. T. (2007). Correlations between academic achievement and anxiety and depression in medical students experiencing integrated curriculum reform. *The Kaohsiung Journal of Medical Sciences*, 23 (8), 379-386.

Üniversite Öğrencilerinin Depresyonlarının İncelenmesi

(Özet)

Problem Durumu: Son yıllarda depresyon tanısı alan, tedavi gören ve depresyon için ilaç kullanan üniversite öğrencilerinin sayısında bir artış gözlenmektedir. Depresyon üniversite öğrencilerinin karşı karşıya kaldığı duygusal problemlerden en ciddi olanıdır çünkü yetişkinleri ve genç yetişkinleri doğrudan ve derinden etkilemektedir. Çünkü üniversite yılları boyunca gençler yetişkin dünyasına adım atmaya başlamaktadırlar. Bir başka deyişle üniversiteye adım atan genç yetişkinler, farklı bakış açıları, farklı yaşam biçimlerini keşfetmeye başlamakta ve değişik rolleri olduğunu fark etmektedirler. Bir yandan bağımsız olmaya çalışırken, bir yandan da kendilerinden beklenen gelişim görevlerinden olan yakın ilişkiler kurma çabasına girmektedir. Üniversite öğrencileri kendi beklentileri ile diğer insanların (arkadaşları ve aileleri) beklentilerini karşılama konusunda denge kurmaya çalışmaktadırlar. Ayrıca, üniversite öğrencileri kendilerini keşfetmeye çalışmakta, bu kendini keşfetme süreci ise zaman zaman benlik saygısında iniş ve çıkışlar yaşamalarına neden olabilmektedir. Ayrıca üniversite öğrencilerinin değişen çevre ve kalma yeri gibi koşullara uyum sağlaması da gerekmektedir. Bu nedenle, üniversite yaşamı hem olasılıkları hem de olumsuz yaşam koşullarını ve stres faktörlerini beraberinde getirmektedir. Sonuçta üniversite öğrencileri yaşadıkları baskı sonucunda öfke, kaygı ve depresyon gibi bazı psikolojik belirti ve sorunlara daha yatkın hale gelebilmektedirler.

Araştırmanın Amacı: Depresyon üniversite öğrencilerinde oldukça yaygın olduğu ve yaşamlarını olumsuz yönde etkilediği için depresyonun incelenmesi çok büyük önem taşımaktadır. Üniversite öğrencilerinin yaşam koşulları ve gelişim özellikleri dikkate alındığında dış görünümünden memnun olmama, akademik başarısızlık ve kalma yerine uyum sağlayamama değişkenleri depresyon için en önemli risk faktörleri arasında yer almaktadır. Bu nedenle bu araştırma üniversite öğrencilerinin depresyon düzeylerini cinsiyet, dış görünümünden memnuniyet, kalma yeri ve akademik başarı değişkenleri açısından incelemek için gerçekleştirilmiştir. Bu araştırmada, (1) üniversite öğrencilerinin depresyon düzeyleri nasıldır?, (2) üniversite öğrencilerinin depresyon düzeyleri dış görünümünden memnun olma düzeylerine göre anlamlı olarak farklılaşmakta mıdır?, (3) üniversite öğrencilerinin depresyon düzeyleri akademik başarılarına göre anlamlı olarak farklılaşmakta mıdır?, (4) üniversite öğrencilerinin depresyon düzeyleri cinsiyet ve kalma yerine göre anlamlı olarak farklılaşmakta mıdır?, (5) üniversite öğrencilerinin depresyon düzeyleri dış görünümünden memnun olma ve akademik başarı değişkenleri kontrol edildiğinde cinsiyet ve kalma yerine göre anlamlı olarak farklılaşmakta mıdır? sorularına yanıt aranmıştır.

Araştırmanın Yöntemi: Bu çalışma Anadolu Üniversitesi'ne devam eden 440 üniversite öğrencisiyle gerçekleştirilmiştir. Katılımcıların 315'i kadın, 123'ü ise erkektir; 102'si birinci sınıf, 105'i ikinci sınıf, 110'u üçüncü sınıf, 123'ü ise dördüncü sınıftır. Araştırma verileri, Beck Depresyon Envanteri, Beden

Organlarından Memnuniyet Ölçeği ve araştırmacılar tarafından bu araştırma için geliştirilmiş olan Kişisel Bilgi Anketi ile toplanmıştır. Kişisel Bilgi Formu üniversite öğrencilerinin cinsiyet, sınıf düzeyi, akademik ortalama ve kalma yeri ile ilgili sorulardan oluşmaktaydı. Data analizinde ise bağımsız gruplar için t testi, iki yönlü varyans analizi ve iki yönlü ANCOVA (Kovaryans) analizi kullanılmıştır.

Araştırmanın Bulguları: Üniversite öğrencilerinin depresyon düzeylerine ilişkin dağılımın ortalaması 12.63, standart sapması ise 8.21 olarak bulunmuştur. Üniversite öğrencilerinin %25'inin depresyon puanları yedi ve yediden düşük, %50'sinin puanları ise 11'e eşit veya 11'den düşüktür, %75'inin puanları ise 17 ve daha altında yer almaktadır. Bu sonuçlara göre, üniversite öğrencilerinin büyük çoğunluğunun depresyon düzeylerinin yüksek olmadığı ifade edilebilir. Araştırmanın bulguları dış görünümünden memnun olmayan üniversite öğrencilerinin depresyon düzeylerinin dış görünümünden memnun olanların depresyon düzeylerinden daha yüksek olduğunu ortaya koymuştur. Ayrıca akademik başarısızlık yaşayan öğrencilerin depresyon yaşama olasılıkları akademik olarak başarılı olanlara göre daha yüksek bulunmuştur. Araştırmanın bulguları üniversite öğrencilerinin depresyon düzeylerinin cinsiyete, kalma yerine ve bu iki değişkenin etkileşimine göre anlamlı olarak farklılaşmadığını ortaya koymaktadır. Kovaryans analiz sonuçları ise üniversite öğrencilerinin depresyon düzeylerinin dış görünümünden memnun olma ve akademik başarı değişkeni kontrol edildiğinde de cinsiyet, kalma yerine ve iki değişkenin etkileşimine göre anlamlı olarak farklılaşmadığını göstermektedir.

Araştırmanın Sonuçları ve Öneriler: Araştırmanın bulguları üniversite öğrencilerinin kalma yeri ve cinsiyetlerinin depresyon düzeylerinde farklılık oluşturmada önemli faktörler olmadığını ortaya koymuştur. Bu faktörler, üniversite öğrencilerinin dış görünümünden memnun olma ve akademik başarı durumları kontrol edildiği halde anlamlı bir farklılığa yol açmamıştır. Öte yandan, üniversite öğrencilerinin dış görünümünden memnun olmaları ve akademik başarıları depresyon düzeylerinde anlamlı etkilere sahiptir.

Depresyon üniversite öğrencilerinin karşı karşıya kaldığı en önemli psikolojik problemlerden birisi olduğu için üniversite öğrencilerine etkili psikolojik yardım hizmetleri sunulması gerekmektedir. Bu nedenle üniversite öğrencilerinin kendini kabul düzeylerini yükselten, iletişim becerilerini ve sosyal ilişkilerini geliştiren farklı depresyonla baş etme programları geliştirilebilir. Bu programların tasarlanıp hazırlanmasında ise dış görünümünden memnun olup olmama ve akademik konular da dikkate alınmalıdır. Bu çalışmanın bazı sınırlılıkları vardır bu nedenle bu araştırma başka çalışmalarla tekrarlanabilir. Bu araştırmanın katılımcılarının sayısındaki sınırlılıktan dolayı, üniversite öğrencilerinin depresyon düzeyleri ile cinsiyet, dış görünümünden memnun olma ve kalma yeri arasındaki etkileşimler daha büyük çalışma gruplarıyla gerçekleştirilebilir. Üniversite öğrencilerinin depresyonlarını daha ayrıntılı olarak ele almak için nitel çalışmalar yapılabilir.

Anahtar Sözcükler: Akademik başarı, kalma yeri, beden doyumu, depresyon, cinsiyet, üniversite öğrencisi.

An Application of Digital Portfolio with the Peer, Self and Instructor Assessments in Art Education

Ayhan Dikici*

Suggested Citation:

Dikici, A. (2009). An application of digital portfolio with the peer, self and instructor assessments in art education. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 91-108.

Abstract

Problem statement: In art education, one of the most difficult tasks is to evaluate the artistic works of students. Portfolio assessment is a method of assessment that is commonly used in artistic education. The major research question investigated in the study is: "What are the advantages and disadvantages of digital portfolio in an art classroom?"

Purpose of the study: The first aim of this study is to investigate digital portfolio assessment in higher art education with the combination of self, peer and instructor ratings. The second aim is to determine opinions and perspectives of the students on the self and on peer and instructor assessment.

Research method: This research contains data collected from 34 students. The students were 14 males and 20 females. Age levels of the students ranged from 21 to 28. The students were asked to prepare their portfolio in the computer environment. The students were given 4 weeks to prepare the digital portfolios. The digital portfolios were evaluated as to the rubric which was previously prepared and adopted by the students and researcher. The students first evaluated their own portfolios and then those of their peers. Finally, the researcher evaluated all of the portfolios.

Results and discussion: Regarding the evaluation of the results, the lowest correlation values were found between the instructor and the peer, and the highest correlation values were found between self and the peer. In the interviews with the students, the students judged the peer assessment to be less trustworthy than instructor assessment. The hesitation of the students about self assessment and the peer assessment might be caused by Turkish culture, because an instructor as a superior authority figure is always considered the most trustworthy person by the Turkish students.

* Asst. Prof. Dr., Nigde University Faculty of Education, adikici@nigde.edu.tr

Recommendations: In terms of educational applications, digital portfolio as an assessment tool should be improved and widely applied in classroom settings where the subjects especially deal with artistic art works. In art teacher training programs, the importance of triadic assessment should be seriously taken into account. This would be vitally important to teach how viable criteria can be set up for the assessments.

Key words: Art education, portfolio assessment, digital portfolio, assessment rubric.

Introduction

Portfolios are used in many fields such as fine arts, marketing, architecture, and education. A portfolio is defined as “a purposeful collection of student works that reflect the efforts, development and successes of the learner” (Jacobson, Sleicher & Maureen, 1999; Mullin, 1998; Paulson, Paulson & Meyer, 1991, p. 60; Stiggins, 1997). Portfolio assessment procedures are increasingly recommended to document students’ performance in school and in extra-curricular activities. During the past two decades, the use of portfolios to assess creativity and giftedness has been widely accepted due to the development of explicit, well-defined procedures for their use (Johnsen & Ryser, 1997). Examples of the students’ work are typically reviewed by knowledgeable individuals (Davis, 1997). Authentic measures such as portfolios, work samples, and biographical information are essential components in identifying talent, for example, in the visual arts.

Properties of portfolio assessment. The most distinctive property of a portfolio makes a learner both assessor and assessed. In this case, apart from being the object of the assessment, the student is both the partner of the assessed object and the assessment (Wolf, 1991, p. 130). Here, the learner actively participates in the selection of the content and determining the selection of the criteria. Therefore, portfolios serve both for teachers and for students. They not only provide an opportunity for students to project their successes, but also provide an opportunity for teachers to evaluate the development and success of their students. Students test their own works and project them on their targets for the future (Stiggins, 1997). Traditional tests do not reveal the development and all the skills of the individual (Barton & Collins, 1993).

Benefits of portfolio assessment. One of the many benefits of a portfolio brings clarity to the fairness in assessing students’ performances. By means of portfolio assessment, assessment is not a secret method. Students judge the quality of their own works and develop standards. In determining the assessment criteria, student-student negotiation and student-teacher negotiation play important roles. In order to shape the assessment criteria, students can be allowed to have a class discussion. This method provides an educational environment both for the students and the teachers (Mullin, 1998). This environment enables the students to be responsible for their own development because it contributes to the activation of the meta-cognitive awareness to monitor their own learning. So, a portfolio assessment is not simply a tool to use to evaluate the end product. It also monitors students’ learning processes. However, it is important to note that an assessment rubric should be reliable and applicable as independently as possible. Students should reach a common decision

by acting together to determine the structure, contents and criteria of the portfolio and the necessary documents under the guidance of their instructor.

Triadic assessment. Some form of self and peer assessments in higher education have been documented in the literature (McConnel, 2002). Gale, Martin and McQueen (2002) developed a triadic (*self, peer and tutor*) assessment by using questionnaire and semi-structured interview techniques at the University of Plymouth in the UK. The research sample was drawn from groups of students studying at different stages of the Bachelor Education and Training program within the Faculty of Arts & Education. Questions of power and knowledge were closely examined, alongside a close critical analysis of the way in which the professional identity and assessment practice style of those involved in the research were influenced by the way they were situated within a variety of different practice contexts. However, researchers suggested that triadic assessment was not a transparent activity to easily generalize due to its lack of universal features and characteristics.

A similar research was conducted by Centra (1994). Centra examined the correlation between scores given by two peer faculty members and four deans to the same portfolios. In the portfolios, faculty documented their accomplishments and wrote personal statements in four areas: teaching effectiveness, service to the college and community, personal credentials, and professional activities. Results of Centra's study indicated that all groups of raters judged the portfolios with very high scores. The deans gave the lowest ratings for total teaching effectiveness. Centra realized that a problem with the portfolio assessment in his study is that no standard criteria was used to rate the portfolios. Centra suggested that when using portfolios for evaluation, a standard criteria by which they are to be judged should be established.

Usage of Portfolio Assessment in Art Education

Eisner (1985) developed a holistic approach, which can deal with both processes and products, to characterize the ways that professional assessors perform tasks similar to those of connoisseurs and critics in the arts. Such a holistic assessment approach is certainly possible for much educational judgment, but it must go beyond the traditional assertion.

According to Madeja (2004), appropriately trained art teachers are capable of judging the aesthetic merits of artwork by using well established techniques in the field. However, standardized tests should not be the only measure, especially in arts assessment. In art education, a portfolio assessment is different from the traditional assessments in terms of scoring. Multiple parties such as the instructor, artists and students could participate in rating a portfolio, but it is important to have an agreement in scoring when the parties rate the portfolios independently. Sabol (2006) conducted a study related to portfolio assessment rated by the art teachers, art students, and artists. Sabol has found impressive results in his study. Fifty-nine elementary, middle, and secondary art teachers, 472 students, and 50 artists were involved in this research. Even though the criteria used to assess the portfolios suggest different priorities for the art teachers, art students, and artists in his study, high levels of agreement were found between the art teachers and students in scoring

the portfolios. Sabol interpreted this as a possible result from art teacher imposed criteria for artwork (p. 10).

The use of digital portfolio assessment in art education. The use of computers in art education is increasing day by day. While art teachers do not doubt that digital portfolios have advantages, they are skeptical about the reliability of using digital portfolios in order to evaluate students' artwork. They want to know whether evaluations of digital reproductions of artwork would be comparable to evaluations of actual artwork. Furthermore, they are concerned as to whether evaluations of portfolios by art educators would be comparable and reliable (Dorn, Madeja & Sabol, 2004).

In a recent study, Dorn and Sabol (2006) carried out a research involving 178 students of 29 secondary school teachers from four school districts in Florida and Indiana by using a digital portfolio. The results suggested that electronic portfolios could be used to reliably estimate students' art performances and scores, because art teachers' evaluations of the digital copies of the actual artwork were consistent with the evaluations of these same works in the actual form.

The Purpose of the Study and Research Question

The literature on digital portfolio assessment suggests that more research is needed that examines the effectiveness of this type of assessment and that reveals its advantages and disadvantages. Therefore, the main aim of this study is to investigate digital portfolio assessment in higher art education with the combination of self, peer and instructor ratings. In addition, the correlation among ratings of the self, peer and instructor on digital artwork is an interest of this study to judge the reliability of this technique. The major research question investigated in the study is: What are the advantages and disadvantages of digital portfolio in an art classroom?

Research Method

Participants

The research involved the fourth year (senior) students in the art teacher training program at a middle-sized Turkish University. The program aims to train art teachers for primary and secondary school levels with four-years of study. Turkish art teachers are trained according to the "Pre-service Art Teacher Training Program (ATTP)" by Kırıçoğlu and Stokrocki (1997). In the ATTP, the first three years mainly focus on developing trainees' art skills, such as art history, media applications, aesthetics, art criticism, computer operations, and drawing-painting skills. Starting with the third year, teacher training courses (i.e., classroom management, teaching methods, and lesson plans) are provided. The last year of the program emphasizes the practicum approach in school settings; where trainees are taken to schools to teach lessons and understand the daily routines in school settings.

The study sample consisted of 34 senior students at an art teacher training program at a College of Education, during the spring semester of the 2005-2006 academic year. The participants in the study were 14 males (41.1%) and 20 females (58.8%). Age levels of the participants ranged from 21 to 28. The average age for students was 21.71 years old.

Data Gathering Method

Both qualitative and quantitative data were utilized in the study. The following sections will clarify the sources of the data and their analysis.

I. Quantitative data. In this study, the students and the researcher utilized the same assessment rubric in order to assess the digital portfolio materials of the students. The criteria for the selected rubric were adopted from http://www.bcpl.net/~sullivan/modules/tips/rubrics_sec/scrapbook.html (Table 1.). Validity of a portfolio assessment is supported in two ways: (1) a well-established appraisal system, which encourages self-development, provides the opportunity for creative problem solving, and (2) this system focuses on individual effort for accepted program goals (Bursch, 1997). Therefore, the basic components of the student performance required for scrapbook preparing were: "content, theme, language & conventions, and overall effectiveness." Point assignment was used for each component of the assessment rubric. Points scale ranged from "two" to "eight." This kind of point assignment provides flexibility for the assessors in scoring (Custer, 1996; Herman, Gearhart & Baker, 1993; Moscal, 2000; Sabol, 2006). The scoring criteria were as follows: Minimal 2 point, Basic 4 point, Good 6 point, and Perfect 8 point.

Table 1
Scrapbook Rubric*

	Perfect 8	Good 6	Basic 4	Minimal 2
Content	Information is complete and is enhanced by accurate and appropriate details. Pictures, photographs, drawings, diagrams, graphs, or other similar devices add to the overall effectiveness of the scrapbook; captions are relevant and explanatory. Space, shapes, textures, and colors provide information themselves and add to the overall effectiveness of the scrapbook. All sources are properly and thoroughly cited; the number/types of sources are exceeded.	Main points are covered but lack some detail. Number and types of visuals are adequate as are captions. Design elements and principles used are adequate. The minimum number/types of sources are present and are cited properly.	Some main points and details are missing. More and better visuals could be used; captions only identify and label rather than explain. Shows evidence of use of some design elements and principles. Sufficient number/types of sources are lacking; citations are not all formatted properly.	Main points are not complete and are greatly lacking in detail. Very little pictorial representation is present; captions are incomplete. There is no consideration of design elements and principles. The list of sources if inadequate in number/types and format of items.

Theme	There is wholeness about the scrapbook; the theme is consistent throughout. The cover clearly identifies the theme.	Most of the information relates to the theme of the scrapbook. The cover is relevant to the contents.	Only a portion of the information relates to the theme of the scrapbook. The cover is unclear in its message.	Confusing and/or inconsistent information.
Language & Conventions	Conventions of spelling, punctuation, and grammar are used with a high degree of accuracy.	Most conventions of spelling, punctuation, and grammar are used accurately.	Common conventions of language are used with some accuracy but there are some mistakes.	A significant number of errors are made in spelling, punctuation, and grammar.
Overall Effectiveness	The requirements of the assignment have been exceeded. The scrapbook is very creative and interesting.	All the requirements of the assignment have been fulfilled. The scrapbook is neat and presentable.	Only some of the assignment requirements are fulfilled. Areas of the scrapbook lack neatness.	Few of the assignment requirements have been met. The presentation as a whole lacks neatness.

*The rubric adopted from <http://www.bcpl.net/~sullivan/modules/tips/rubricssec/scrapbook.html>

The process of preparing the digital portfolios was completed as follows: First of all, students were required to identify the aims for their portfolios they would prepare. In the first week, it was carried out into the identification of the aims. The significance and necessity of the aims, this frame, which would be provided for the portfolios the students would improve, were explained to the students.

Identification of the necessary features for a qualified scrapbook and the presentation of the assessment rubric were in the frame. The features that would be necessary for a qualified scrapbook were determined by face-to-face discussions with students. The assessment rubric which had been improved beforehand was presented in the frame features. These frame features highlight the major components for a qualified scrapbook that would be carried out by students. Also, the duties and the rubric indicating the criteria for evaluation and scoring were explained to the students in detail.

The scrapbook examples and the rubric were sent to the e-mail addresses of the students. The paper-based samples and the rubric were given to the students in the classroom. Thus, the students had an opportunity to know not only which samples of the work were good, average and unsuitable but also for what reasons those works were evaluated in that way. They were given the chance to improve their works in that frame.

Each student was required to prepare at least four scrapbooks for his/her portfolio that s/he would hand in at the end of the research. Each scrapbook contains a well-known artist's paintings and bibliography. In the instruction, it was pointed

out that the selected artists in the scrapbooks should be from different time periods, from the renaissance era to the present-day. Students also had freedom to prepare additional scrapbooks. Students had 4 weeks to turn in their scrapbooks. With 4 hours in a computer-based environment for the students, a total of 8 hourly practices were carried out each week. During the research session, the students made use of various web pages and software package programs such as Word, Paint Brush, Power Point Presentation and Illustrator, and Photoshop to prepare their digital portfolios.

This course was not limited to classroom practices and face to face discussions. Discussions were also taking place on the online environment to extend their learning. Although we did not use an online discussion board, students asked their questions, presented their works, and shared their ideas with their classmates via electronic mail and msn messenger. In this way, each student had an important opportunity to hear the critiques of the others for their works. In each stage of the works (at the end of each course) the quality of the students' works were evaluated. In addition, necessary corrections and feedback were provided in order that there was a scope for students to take note of incorrect and inefficient parts of their works and to improve on them.

The explanation for the storage of the students' works and how they would present prepared portfolios was given in the last week. There were some alternatives for the students to store their works on computers, in electronic mail addresses, or on CDs, DVDs or flash disks. Students were permitted to use a scanner in order to transfer the works done by pen and paper to the digital environment. Students presented the drafts of each of their own works in a chronological order.

The students (self and peer) and the researcher participated in the assessment. Each student first evaluated his/her own portfolio according to the rubric and then each student evaluated the portfolio of his/her friend according to the rubric. In peer assessment, the students did not know whose portfolio they were evaluating (blind review). Finally, the researcher evaluated all portfolios. The researcher has expertise in art education and the use of portfolio assessment in educational settings. The assessors made individual assessments without external influences. In addition, the assessors didn't do a different evaluation for digital portfolios. In other words, the assessors made an assessment with respect to the assessment rubric which was developed for students' digital portfolios.

Analysis of quantitative data: After collecting the scores given by the three parties--self, peer and the researcher--at the end of the research, Pearson correlation coefficients were calculated by using the SPSS 13.0 statistical package program. The significance levels of the coefficient values were determined according to Hopkins (1997). The coefficient values of 0.10-0.30, 0.30-0.50, and 0.50-0.70 are considered low, medium, and high correlations, respectively.

II. Qualitative data. After the evaluation procedures were completed for the portfolios, several open-ended and semi-structured questions were asked to the students. The examples from the questions sets are as follows: What do you think

about that evaluation procedure overall? What kind of software did you prefer when you prepared your portfolio? Why? To what extent were you confident as you evaluated your own portfolio? Why? To what extent were you confident as you evaluated the portfolio of one of your friends? Why? What do you think about this triple evaluation? What are your overall experiences from all of these? The interviews were audio-taped and then transcribed. In the analysis, after the data was encoded, the themes were found and then arranged (Strauss & Corbin, 1998). Finally, each theme was expressed with frequencies and each was interpreted.

Results and Discussion

Evaluation of Statistical Findings

In the correlation analyses, while the lowest correlation was .19 between the researcher and peer in scoring the *Language & Conventions*, the highest correlation was found to be 0.50 between self and peer ratters in scoring the *Overall Effectiveness* and in scoring *Theme* (See Table 2). Compared to the ranges of values determined by Hopkins (1997), the values in Table 2 can be accepted as at the medium level. Also, Cohen and Lea (2003) stated that the value at the level of 0.30 was a medium level value for social sciences.

Table 2

The Correlations among the Assessors (R: Researcher, S: Self, P: Peer)

Sub-Components	Assessors (R,S,P)			
	R	S	P	
Content	R	1.000		
	S	0.491*	1.000	
	P	0.420*	0.334*	1.000
Theme	R	1.000		
	S	0.502*	1.000	
	P	0.389*	0.323**	1.000
Language & Conventions	R	1.000		
	S	0.462*	1.000	
	P	0.193**	0.499*	1.000
Overall Effectiveness	R	1.000		
	S	0.467*	1.000	
	P	0.372*	0.504*	1.000

*p< 0.01, **p< 0.05 significance

It was found that the highest correlation was between peer and self scoring (0.504) and the lowest correlation was between the researcher and peer (0.193). However, Sabol (2006) found greater levels of agreement in rankings which were found amongst the art teachers and students. Sabol interpreted this as possibly being the result of art teacher imposed criteria for artwork. The art teachers must be concerned with developing a range of knowledge and skills among their students.

According to Sabol, students in school focus on what art teachers teach. Students at home focus on skills with media and personal satisfaction derived from their art.

Pitts, Coles and Thomas (1999) found the reliability of individual assessor's judgements (i.e., their consistency) as moderate, but inter-rater reliability did not reach a level that could support making a safe summative judgment. They said, "What can we learn from qualitative approaches? Qualitative research takes an interpretive, naturalistic approach to its subject matter; qualitative researchers study things in their natural settings, attempting to make sense, or interpret, phenomena in terms of the meanings that people bring to them" (p.519). In another research, Pitts, Coles, Thomas and Smith (2002) found an agreement between assessors ranged from "slight" to "fair." A kappa coefficient of 0.50 reflects the moderate agreement between assessors.

Evaluation of Interviews and Observations

From a grounded theory approach to the analysis of the data (Strauss & Corbin, 1998), five broad analytic categories were built under which the experiences and opinions of students can be considered. The categories were related to the preparation of the digital portfolios and assessment processes. The features used in forming scrapbooks in the digital portfolio were taken into consideration in the categorization. The frequencies and percentages of the observed themes among students are presented in Table 3. The categorizations of the themes were based on the analyses of the interviews and observations of the researcher about the preparation and evaluation phases of the digital portfolios. The sums of the percentage in the Table 3 are not 100% because a majority of the students meet the criteria for more than one category.

Table 3

Frequencies and Percentages of the Categories

Categories	f	%
Approval of the researcher assessment	29	85.2
Paucity of web pages	28	82.3
Hesitation of the peer assessment	25	73.5
Difficulty of computer software	23	67.6
Hesitation of the self assessment	21	61.7

Approval of the researcher assessment. For students, it sounds normal if the only assessor is the trainer. 85.2% of the students thought in that way but they did not have such a confident for themselves and their peers. This might be because they saw themselves as inexperienced in assessment. Some students such as student A, a male student, expressed that if the trainer (researcher) also participated in the assessment, then it was okay.

Student A- *"I was happy when I found out three persons will grade one portfolio. So, you will also give points, too...Then, there is no problem..."*

In fact, self assessment and peer assessment were very strange to the students because they had gotten used to trainer or instructor assessment over the years. Thus, their hesitations and avoidances for the self and the peer assessments can be considered as a normal consequence.

Paucity of web pages. The students stated that they easily obtained the paintings of the artists from the web pages; on the other hand, they didn't find enough information about the artists and their paintings. 82.3% of the students complained about the insufficiency of the web pages. However, some students like student B weren't very constrained in developing content since they prepared the scrapbooks with the paintings from the art magazines.

Student B- *"I could not find detailed information about Paul Gauguin on the web pages. There were a lot of samples from his works, but there were little information related to his pictures. So, I used the art magazines and scanned his pictures and transferred them to the digital environment. Hope you have not got angry with me!..."*

Hesitation of the peer assessment. The students first objected when they were asked to assess the portfolios of their friends. It was observed that they did not want to take responsibility. They were told they would be teachers in the future and they would have to assess and grade their students' works. 73.5% of the students thought that peer assessment might be distrustful. Similar findings were found in a research of Gale et al. (2002). The students hesitated to evaluate their friends' portfolios. For example, student C, a female student, did not think that she could be objective in assessing someone's work.

Student C- *"Sir, you want me to assess my friend's portfolio!...Shall I know whose portfolio I will be assessing?. I am not sure if I will be able to evaluate it objectively."*

With this in mind, student C has in fact accepted that one of the dimensions of the assessment is objectivity. Although she knows the criteria for the assessment, she has low self-confidence.

Difficulty of computer software. The classroom observations let the researcher identify some significant problems. The students first tried to learn how to use computer software that they need for preparing a scrapbook. The four weeks given to the students for their assignment was not a long period of time, because some students apparently also needed to develop specific skills to use the necessary software. Therefore, specific software skills needed to be introduced and practiced.

The software popular among the students to prepare the digital portfolios were *Paint Brush, Microsoft PowerPoint, and Word Processor*. These softwares were seen as somewhat familiar and generally available on all computers. However, it was difficult for the students to use some package programs like *Illustrator and Photoshop*. Such difficulties were expressed by 67.6% of the students. Also,

observations of the researcher in the classroom supported such difficulties experienced by the students. An interview segment of student D, a female student, exemplifies how students had a hard time with the software.

Student D- *“At the beginning, the idea of preparing a scrapbook seemed great to me; however, when I sat at the computer, the first thing that came to my mind was to make a search over the internet. But I could not use a software by which I am able to prepare my work. It is very difficult to use the Photoshop software which you have recommended, so I preferred to use an easy software.”*

It was noticed that there were differences between male and female students regarding the difficulty that they experienced in using the software. While 90.0% of the female students expressed some difficulty in using the software programs, only 35.7% of the male students expressed similar difficulties in using the software. Female students frequently preferred to use Paint Brush, Power Point and Word Processor. Male students showed more effort in using the computer software. Male students used some of the software that female students used too; however, some male students were able to figure out how to use some difficult software like Photoshop. Some male students used Photoshop to arrange the dpi resolutions and sizes of the pictures that they found from the internet. For example, student E, a male student, expressed that although Photoshop is professional software, it is not necessary to be competent in it for their assignments.

Student E- *“When you said we can use the Photoshop software, I was astonished. As far as I know, it is a professional software program; however, all I needed to enlarge and reduce the sizes of the pictures. It was difficult at the beginning but I think I have overcome it...”*

Student E said that Photoshop was professional software; however, he figured out simple functions of Photoshop to prepare his own work. Although he found it difficult to use this software at the beginning, he overcame that problem; however, female students made no attempt to use the Photoshop software.

Hesitation of the self assessment. The researcher received interesting responses when students were asked their opinions about their own works. As the students were only familiar with the traditional evaluation, it was very interesting for them to evaluate their own works. 61.7% of the students thought that their evaluation might be wrong so it would be more suitable if I do the evaluation. A female student, student G, gave an interesting response when the researcher asked how she felt about the self assessment.

Student G- *“Sir, why do you have us make the evaluations?...I don't believe I can evaluate my own study in a truly way even if it is done according to the rubric! You give us too much power. Keep some of them in your own hands...”*

As you notice, Student G felt that she was given too much power for the evaluation because she thought that the evaluations should be under the heavy control of the instructor. Her thoughts might be an indication that she is very loyal to

the authority and that she accepts authority without questioning it at all; however, most students had already had this thought, too, before the start of self assessment. However, their worries and hesitations disappeared after the self assessment, because they had an opportunity to judge their own study. Mullin (1998) says that the most important benefit of portfolio assessment is the opportunity for the student to assess his own work.

A Theoretical View to the Findings

On the basis of the presented data, a model was generated to determine the students' perspectives on digital portfolio assessment (Figure 1).

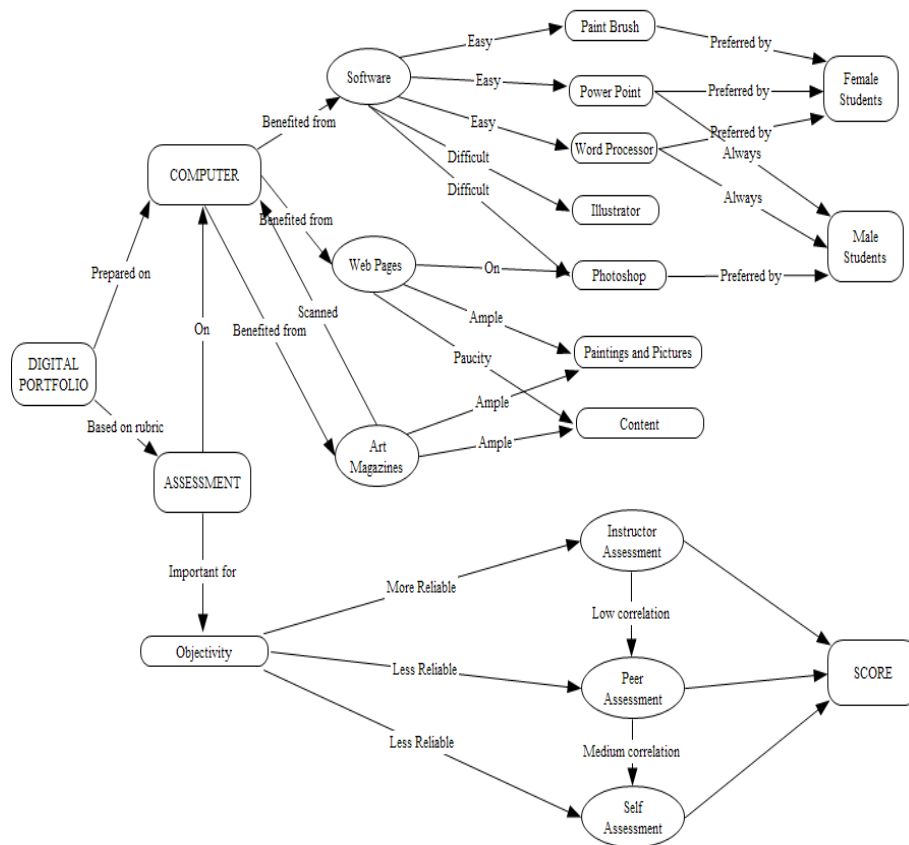


Figure 1. Integration of the findings

At the left of the diagram, there was the DIGITAL PORTFOLIO prepared on the COMPUTER where the ASSESSMENT was made. In the preparation of the digital portfolios, the students benefited from the SOFTWARE and WEB PAGES. Based on

the data taken from the students, the software fall into two categories: difficult and easy. The easy ones were PAINT BRUSH, POWER POINT and WORD PROCESSOR. The difficult software included ILLUSTRATOR and PHOTOSHOP. While Paint Brush, Power Point and Word Processor were preferred by FEMALE STUDENTS, Paint Brush and Power Point were constantly preferred by MALE STUDENTS. Photoshop, which was difficult software to use, was used by some male students. No student used Illustrator. The students benefited from the web pages in the creation of digital portfolios. While students found many PAINTINGS and PICTURES on web pages, they were not enough in terms of CONTENT. Some male students transferred the pictures they took from web pages to Photoshop. Some students scanned the pictures and contents of ART MAGAZINES and they placed them in the computer environment.

In the second part of the diagram, there was the evaluation of the digital portfolios on the COMPUTER, ASSESSMENT. This assessment was made according to the rubric. OBJECTIVITY was important in assessment. In this assessment, INSTRUCTOR ASSESSMENT was more important. SELF and PEER ASSESSMENT was more trustworthy. There was low correlation between instructor assessment and peer assessment, and medium correlation between peer assessment and self assessment. The average values of all three assessments determine the score of the digital portfolio.

Conclusion

It is important to prepare a rubric in the beginning in terms of harmony amongst self, peer and the researcher's points. A lower correlation coefficient, reflecting less than moderate agreement, is unlikely to be acceptable. According to Pitts et al. (2002), whether a coefficient of 0.50 is acceptable is open to debate. Certainly, the higher value is the better. Nevertheless, these results show that discussion between assessors increases reliability above the levels often achieved in assessments of professional competence, and is an improvement on individual assessment. This result is also a result recommended by Centra (1994); however, it was determined the score points in the rubric ranged from 2 to 8. If it was arranged in way that the score points in the rubric ranged from 1 to 4, correlation values amongst the pointers might have been higher. The hesitation of the students about self assessment and peer assessment might be caused by Turkish culture, because an instructor as a superior authority figure is always considered the most trustworthy person by the Turkish students. Korkmaz and Kaptan (2003) suggested the practical information to the teachers on portfolio assessment both in their pre-service and in-service education should be given. Even though portfolio assessment has many difficulties, the use of portfolio assessment in our schools and teacher training systems should be taken into consideration. Because the portfolio assessment approach has better features than the

traditional assessment tools with regard to reflecting the efforts, development and successes of the students

The capabilities of the students to justify their own works support their learning. Peer assessment can enable to the students a fair and an honest pointing habit; however, this evaluation in this study is not a generalizable one. Students from different training fields and age groups can reveal different opinions. Gale et al. (2002) stated that “triadic assessment was clearly not a transparent activity that revealed easily generalizable and universalizable features and characteristics.”

As a result of the current study, it is suggested that digital portfolio as an assessment tool should be improved and widely applied in classroom settings where the subjects especially deal with artistic art works. In the art teacher training program, the importance of triadic assessment should be seriously taken into account. It is important to learn how to define assessment criteria. Also, it would be effective for Turkish students to overcome their cultural barrier regarding peer and self assessment types. Furthermore, it is suggested that the assessment exercises employed in this study would contribute to the art teacher candidates’ professional development. In fact, the Turkish Ministry of National Education suggests as assessment tools the self assessment, group assessment, peer assessment, checklist, portfolio assessment and performance assessment in art education (MEB, 2006, p. 100).

Limitations of the Study

This study was conducted in the field of art education. Digital portfolios prepared with scrapbooks by the students were evaluated. Different results could be achieved from the data collected from different cultures’ students. In this study, the students encountered such an assessment for the first time. Different results could be achieved with the students who have already been made familiar with such an evaluation.

References

- Barton, J. & Collins, A. (1993). Portfolios in teacher education. *Journal of Teacher Education*, 44(3), 200-210.
- Bursch, C. B. (1997). Finding out what’s in their heads: Using teaching portfolios to assess English education students and programs. In K. Yancey and I. Weiser, (Eds.). *Situating Portfolios: Four Perspectives*. Logan: Utah State Press.
- Centra, J. A. (1994). The use of the teaching portfolio and student evaluations for summative evaluation. *Journal of Higher Education*, 65(5), 555-570.
- Cohen, B. H. & Lea, R. B. (2003). *Essentials of statistics for the social and behavioral sciences*. New York: John Wiley & Sons.
- Custer, R. L. (1996). Rubrics: An authentic assessment tool for technology education. *The Technology Teacher*, 55(4), 27-37.

- Davis, G. A. (1997). Identifying creative students and measuring creativity. *Handbook of Gifted Education*. (2nd ed.). In N. Colangelo and G. A. Davis (Eds.). Boston: Allyn & Bacon.
- Dorn, C. M. & Sabol, F. R. (2006). The effectiveness and use of digital portfolios for the assessment of art performances in selected secondary schools. *Studies in Art Education*, 47(4), 344-362.
- Dorn, C. M., Madeja, S. S., & Sabol, F. R. (2004). *Assessing expressive learning*. Mahwah, NJ.: Erlbaum.
- Eisner, E. W. (1985). *The art of educational evaluation*. London: Falmer Press.
- Gale, K., Martin, K., & McQueen, G. (2002). Triadic assessment. *Assessment & Evaluation in Higher Education*, 27(6), 557-567.
- Herman, J. L., Gearhart, M., & Baker, E. (1993). Assessing writing portfolios: Issues in the validity and meaning of scores. *Educational Assessment*, 1(3), 201-224.
- Hopkins, W. G. (1997). New view of statistics. Retrieved September 13, 2004 from <http://www.sportsci.org/resource/stats/effectmag.html#cohen>
- Jacobson, W., Sleicher, D., & Maureen B. (1999). Portfolio assessment of intercultural competence. *International Journal of Intercultural Relations*, 23(3), 467-492.
- Johnsen, S. K. & Ryser, G. K. (1997). The validity of portfolios in predicting performance in a gifted program. *Journal for the Education of the Gifted*, 20, 253-267.
- Kırıçoğlu, O. & Stokrocki, M. (1997). İlköğretim sanat öğretimi [Primary school art teaching]. *Öğretmen eğitimi dizisi* [Teacher training series]. Ankara: Ajans-Türk Basın ve Basım A.Ş.
- Korkmaz, H. & Kaptan, F. (2003). İlköğretim fen öğretmenlerinin portfolyoların uygulanabilirliğine yönelik güçlükler hakkındaki algıları [Primary school science teachers' perceptions about barriers toward practicability of portfolios]. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 13(1), 159-165.
- Madeja, S. S. (2004). Alternative assessment strategies for school. *Art Education Policy Review*, 105(5), 3-13.
- McConnell, D. (2002). The experience of collaborative assessment in e-learning. *Studies in Continuing Education*, 24(1), 73-92.
- MEB (2006). İlköğretim sanat etkinlikleri dersi programı ve kılavuzu 1.8 Sınıflar [Curriculum and Guide of the Primary School 1-8 Grade Art Activities Course]. Ankara: Milli Eğitim Basımevi.
- Moscal, B. M. (2000). Scoring rubrics: What, when and how?. *Practical Assessment, Research & Evaluation*, ERIC Clearinghouse, vol.7 no.3, Retrieved April 25, 2005 from the World Wide Web: <http://PAREonline.net/getvn.asp?v=7&n=3>.
- Mullin, J. A. (1998). Portfolios: Purposeful collections of student work. *New Directions for Teaching and Learning*, 74, 79-87.
- Paulson, F. L., Paulson, P. R., & Meyer, C. (1991). What makes a portfolio, a portfolio? *Educational Leadership*, 48(5), 60-63.

- Pitts, J., Coles, C., & Thomas, P. (1999). Educational portfolios in the assessment of general practice trainers: Reliability of assessors. *Medical Education*, 33(7), 515-520.
- Pitts, J., Coles, C., Thomas, P., & Smith, F. (2002). Enhancing reliability in portfolio assessment: Discussions between assessors. *Medical Teacher*, 24(2), 197-201.
- Sabol, F. R. (2006). Identifying exemplary criteria to evaluate studio products in art education. *Art Education*, November, 6-11.
- Stiggins, R. J. (1997). *Student-centred classroom assessment* (2nd ed.). Upper Saddle River, New Jersey: Prentice-Hall, Inc.
- Strauss, A. & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. London: Sage Publication
- Wolf, K. (1991). The school teacher's portfolio: Issue in design, implementation, and evaluation. *Phi Delta Kappan*, 2, 129-136.

Sanat Eğitiminde Akranın, Kendisinin ve Öğreticinin Değerlendirmesi İle Dijital Süreç Dosyasının Bir Uygulaması

(Özet)

*Problem Durumu:*Süreç dosyaları birçok alanda kullanıldığı gibi eğitim alanında da kullanılmaktadır. Süreç dosyası gibi çalışma örnekleri görsel sanatlarda bireyin yeteneğinin ve gelişiminin belirlenmesinde önemli role sahiptir. Süreç dosyasının en ayırt edici özelliği öğreneni hem değerlendiren hem de değerlendirilen yapmasıdır. Öğrenci değerlendirme ölçütlerinin ve içeriğinin belirlenmesine aktif olarak katılır. Böylece, süreç dosyası hem öğrencilere hem de öğretmenlere hizmet eder. Süreç dosyası değerlendirme sanat eğitiminde yaygın bir şekilde kullanılan bir değerlendirme şeklidir. Sanat eğitiminin en zor işlerinden birisi öğrencilerin sanat çalışmalarının değerlendirilmesidir. Bu araştırmanın en önemli araştırma sorusu, bir sanat sınıfında dijital süreç dosyasının avantaj ve dezavantajları nedir?

*Araştırmanın Amacı:*Bu araştırmanın temel amacı eğitim fakültesi resim iş öğretmenliğinde öğrencinin kendisinin, akranının ve öğreticinin puanlama birleşimi ile dijital süreç dosyası değerlendirmeyi araştırmaktır. Araştırmanın bir diğer amacı ise öğrencinin kendisi, akranı ve öğreticinin değerlendirmesi hakkında öğrencilerin bakış açılarını ve düşüncelerini belirlemektir.

*Araştırmanın Yöntemi:*Bu araştırmada, Eğitim Fakültesi Resim İş Öğretmenliği son sınıf öğrencilerinden 34 öğrenci üzerinden veri toplanmıştır. Bu öğrencilerin 14'ü erkek 20'si kız öğrencidir. Öğrencilerin yaşları 21 ile 28 yaş arasında değişmektedir. Öğrencilerden bilgisayar ortamında süreç dosyalarını hazırlamaları istenmiştir. Süreç dosyalarının

içeriğini Rönesanstan günümüze istedikleri sanatçıları tanıtan kesme yapıştırma kitapların (scrapbook) oluşturması söylenmiştir. Öğrencilere dijital süreç dosyalarını hazırlamaları için dört haftalık bir süre verilmiştir. Haftalık 8 saatlik uygulamanın 4 saati bilgisayar ortamında yapılmıştır. Öğrencilerin dijital süreç dosyalarını hazırlamak için çalışmaları süresince Word, Paint Brush, Power Point Presentation, Illustrator ve Photoshop gibi bilgisayar programlarından ve çeşitli internet sayfalarından yararlanabilecekleri söylenmiştir. Öğrencilerin bilgisayar ortamında hazırladıkları çalışmaları elektronik postalarında, CD ROM, DVD, taşınabilir bellek gibi ortamlarda depolanmıştır. Öğrencilerin dergi, gazete gibi kâğıt ortamındaki çalışmaları dijital ortama aktarmaları için tarayıcı (scanner) kullanmalarına izin verilmiştir. Öğrenciler süreç dosyalarının içerisinde yer alan çalışmalarını (scrapbook) tarihsel bir sıra düzenine göre sıralayıp dijital ortamda sunmuşlardır. Öğrencilerin hazırladıkları süreç dosyaları daha önceden belirlenmiş ve hazırlanmış bir yönergeye (rubric) göre değerlendirilmiştir. Öğrenciler ilk olarak yönergeye göre kendi çalışmalarını daha sonra yine aynı şekilde yönergeye göre arkadaşlarının çalışmalarını değerlendirmişlerdir. Son olarak tüm süreç dosyaları araştırmacı tarafından yine aynı yönergeye göre değerlendirilmiştir. Araştırmanın nicel boyutunda, öğrencilerin süreç dosyaları için kendisine ve akranına verdiği puanlar ile araştırmacının verdiği puanların arasındaki ilişkiye (korelasyona) bakılmıştır. Araştırmanın nitel boyutunda ise öğrenciler ile yarı yapılandırılmış görüşme tekniği kullanılmıştır. Öğrencilere çalışma süresi içerisinde yaşadıkları zorluklar ve kolaylıklar ile yaptıkları değerlendirme ile ilgili açık uçlu sorular sorulmuştur. Öğrencilerin kendi çalışmasını, arkadaşının çalışmasını ve öğreticinin onların çalışmalarını değerlendirmeleri hakkında görüşleri alınmıştır. Öğrenciler ile yapılan görüşmeler ses kaydı yapılarak daha sonra metne dönüştürülmüş ve analiz edilmiştir.

Sonuçlar ve Tartışma: Araştırmanın nicel sonuçlarında en düşük puanlama korelasyonu araştırmacının kendisi ile öğrencinin akranı arasında bulunmuştur. En yüksek puanlama korelasyonu ise öğrencinin kendisi ile akranı arasında bulunmuştur. Araştırmanın nitel boyutu için öğrenciler ile yapılan görüşmelerde ise beş temel kategorinin oluştuğu görülmüştür. Bu kategoriler frekans ve yüzdesi en yüksekten en düşüğe doğru şöyledir: Araştırmacının değerlendirmesini onaylama, web sayfalarının yetersizliği, akran değerlendirmede tereddüt, bilgisayar programlarının zorluğu ve kendini değerlendirmede tereddüttür. Öğrencilerin öğreticinin değerlendirmesine daha çok güvendikleri görülmüştür. Öğrenciler internet sayfalarından yeterince yararlanamadıklarını belirtmişlerdir. İnternet sayfalarında hazırlayacakları süreç dosyaları için sanatçıların resimlerinin bol olduğunu ancak sanatçılar ve eserleri hakkında yeterince bilgi bulamadıklarını belirtmişlerdir. Ayrıca öğrenciler kesme yapıştırma kitaplarını oluşturmada kullandıkları bilgisayar programlarında da zorluk yaşamışlardır. Profesyonel bilgisayar programlarında kız öğrencilerin daha

çok zorlandıkları hatta bu programları tercih etmedikleri bulunmuştur. Erkek öğrencilerin bazılarının ise profesyonel programları tercih ettiği görülmüştür. Öğrencilerin arkadaşının süreç dosyasını değerlendirmede tereddüt ve endişeye kapıldıkları öğreticinin değerlendirmesinden daha az güvenilir olacağını belirtmişlerdir. Öğrencilerin kendi süreç dosyasını ve arkadaşının süreç dosyasını değerlendirmede yaşadıkları tereddüdün Türk Kültürü'nden kaynaklandığı düşünülebilir. Çünkü Türk Kültürü'nde öğretmen ya da öğretici otorite, bilgi sahibi ve daima öğrenciler tarafından en güvenilir görülen kişi olarak düşünülür.

Öneriler: Eğitimsel uygulamalar bakımından bir değerlendirme aracı olarak dijital süreç dosyası özellikle sanatsal çalışmalar ile ilgili sınıf ortamlarında yaygın bir şekilde uygulanmalı ve geliştirilmelidir. Sanat öğretmeni eğitim programlarında bu araştırmadaki gibi üçlü değerlendirme yaklaşımına ciddi şekilde yer verilmelidir. Değerlendirmede yönerge çok önemlidir. Değerlendirme yönergesinin kriterleri açık ve net olmalı puanlayıcılar arasında uyumun düşük olmasında neden olmamalıdır.

Anahtar Sözcükler: Sanat eğitimi, süreç dosyası değerlendirme, dijital süreç dosyası, değerlendirme yönergesi

The Effects of Computer Animations and Cooperative Learning Methods in Micro, Macro and Symbolic Level Learning of States of Matter

Kemal Doymuş*
Ümit Şimşek**
Ataman Karaçöp***

Suggested Citation:

Doymus, K., Simsek, U., & Karacop, A. (2009). The effects of computer animations and cooperative learning methods in micro, macro and symbolic level learning of states of matter. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 109-128.

Abstract

Problem Statement: In chemistry education, most first-year undergraduates experience difficulties understanding three related meanings of topics: the macroscopic meaning, the microscopic meaning, and the symbolic meaning. The present study deals with the multiple (macro, micro, and symbolic) meanings of topics in the states of matter in college chemistry.

Purpose of Study: This study investigated the effects of computer animations and cooperative learning on students' comprehension of chemistry topics at the macro, micro, and symbolic levels.

Method: The subjects of the present study comprised 64 university students enrolled in three classes of the general chemistry course taught in the first year of the elementary education science teacher program during the 2006-2007 academic year. One of the classes was used as the cooperative group, the second class was used as the animation group, and the third class served as the control group. Three dependent variables were measured: micro level meaning, non-micro level (macro and symbolic) meaning and students of varying reasoning abilities. Understanding of the micro level was determined by a States of Matter Test (SMT). Understanding of the non-micro level (macro and symbolic) was determined by a Course Exam

* Corresponding Author. Assoc. Prof. Dr., Atatürk University Kazım Karabekir Faculty of Education, kdoymus@atauni.edu.tr

** PhD, Atatürk University Kazım Karabekir Faculty of Education, simsekum@atauni.edu.tr

*** Atatürk University Kazım Karabekir Faculty of Education, akaracop@atauni.edu.tr

(CE). Students' varying reasoning abilities were determined by a Test of Logical Thinking (TOLT).

Findings and Results: A one-way analysis of variance (ANOVA) found no significant differences in the TOLT scores of the control, animation, and cooperative groups. Results showed significant differences in the SMT scores by treatment even when the effects of the TOLT scores were removed. The control group had SMT scores that were significantly different from those of the animation and cooperative groups, while the scores of the two experimental groups did not differ significantly. In addition, no differences were found in terms of understanding at the non-micro level (macro and symbolic). Furthermore, for micro level meanings depicted in this study, the maximum effects were achieved with both computer animations and cooperative learning methods.

Conclusions and Recommendations: This study demonstrated that cooperative learning methods used in small groups helped students understand the states of matter at the micro level and develop cognitive models. Consequently, when correct and suitable learning methods are used, we think that it is more likely that the sources of misunderstanding will be remedied, while at the same time micro level understanding will improve.

Keywords: Computer animation, cooperative learning, micro level, state of matter

Acknowledgement: Parts of this paper were presented at the 1st National Chemistry Education Conference, 20–22 June 2007, Istanbul.

In general science and chemistry education, three levels of understanding have been identified (Johnstone, 1991; Gabel, 1993): (1) the macroscopic level, that is, the meanings of topics are expressed in terms of phenomena, substances, energy, and so on; (2) the microscopic (sometimes called submicroscopic) level, that is, the meanings of topics are expressed in terms of molecules, atoms, ions, and so on; and (3) the symbolic level, that is, the meanings of topics are expressed in terms of formulae, equations, ionic drawings, and so on.

The present study deals with the multiple [micro and non-micro (macro and symbolic)] meanings of topics in the states of matter in college chemistry. College chemistry students have experience going from the macro meaning to the micro meaning and back to the macro, using formula, atomic symbols, and so on. For them, the formula H_2O refers to the substance of water as well as to a single molecule of water. Their mental switching between macro and micro aspects of science curriculum topics is conducted easily and almost automatically (Johnstone, 1993). However, students often experience difficulties understanding the multiple meanings of chemistry topics. For example, they have to learn to consider the phase changes of substances, but they also have to learn the microscopic meaning in terms of the rearrangement of particles, and the symbolic representation in terms of chemical equations (words, iconic drawings, formulae). For them, the conceptual demands of shifting between the three meaning domains can be overwhelming.

Their difficulties in understanding macro-micro meanings—for instance, explaining observations of phenomena in terms of interactions between particles—have been reported in several studies (Benson, Wittrock, & Bauer, 1993; Treagust, Chittleborough, & Mamiala, 2003). Students also appear to experience difficulties understanding symbolic representations; for instance, they handle symbols as algebraic entities without clear interpretations from a macro-micro perspective (Ben-Zvi, Eylon, & Silberstein, 1988; Friedel & Maloney, 1992). As learning to link phenomena with particles and symbolic representations constitutes one of the most important objectives of science education, chemistry students need to develop the ability to teach the macro, micro, and symbolic issue. We assume that students, being educated as chemists, have developed the habit of discussing multiple meanings in a flexible and implicit manner. As a consequence, students encounter difficulties when learning topics at a macro-micro-symbolic interface. To eliminate difficulties in understanding the multiple meanings of general chemistry topics, new educational methods have developed as alternatives to teacher-centered lecturing, such as animation, hands-on and cooperative learning, and inquiry-based and student-centered instruction.

Animation models are useful for teaching chemistry, general science, physics and biology at the micro level. Two-dimensional animated computer models show the dynamic characteristic of chemistry (Sanger, Phelps, & Fienhold, 2000; Ebenezer, 2001). Animated models in three dimensions can be rotated and used to teach spatial relationships (Barnea, & Dori, 1996; Fleming, Hart, & Savage, 2000). Animations are also effective aids for teaching micro and non-micro meanings that involve motion at the molecular level (Gilbert, 2005). An animation is considered three dimensional if it has foreshortened lines, overlapping lines, differences in the relative sizes of objects, and distortion of angles (Shubbar, 1990). Animation models can be viewed by students at a computer terminal, on their own time and as many times as desired, or in the classroom by projection (Theall, 2003). When learning with molecular level representations, students construct mental models based on their observations that are personal, qualitative, and often incomplete, because they often do not understand the underlying concepts that the model represents (Greca, & Moreira, 2000).

Students must learn to navigate model types to solve problems like chemists do (Barnea, & Dori, 1996). Studies show that students who use a combination of model types representing the same concept have a better understanding of molecular level chemistry (Wu, Krajcik, & Soloway, 2001). Sanger and Badger (2001) found that students who viewed electron plots and animations as a supplement to the traditional wooden molecular model kits and demonstrations to learn about molecular polarity and miscibility responded correctly more often on hourly exams than did students who did not view the electron plots and animations. In addition to the availability of computer models, simulations have been created that mimic reactions or systems in areas of science that are too dangerous or costly for students to study otherwise (Yair, Mintz, & Litvak, 2001). Moreover, several chemistry education researchers have demonstrated that computer animations can help students think about chemical processes on the molecular level (Williamson, & Abraham, 1995).

Cooperative learning can be defined as a learning method where students combine in small mixed groups and help each other for a common academic aim, boost each other's self-esteem, develop communication abilities, increase problem solving and critical thinking abilities, and take an active part in learning (Eilks, 2005; Gillies, 2006; Hennessy & Evans, 2006; Lin, 2006). Cooperative learning, which is more efficient than other instructional methods, is widely used in education (Siegel, 2005). In recent years, it has become a strong alternative to the standard education strategies used in high schools and universities. The reason for its popularity is the fact that it gives students the chance to learn from each other's different approaches and decisions by cooperating according to the strategies and problem-solving techniques that are used (Bearison, Magzomes, & Filardo, 1986).

Studies related to cooperative learning in chemistry education show that, on average, using aspects of cooperative learning can enhance chemistry achievement for high school and college students. Based on the results of these studies, it is strongly recommended that chemistry instructors continue incorporating cooperative learning practices into their classes (Bowen, 2000; Doymuş, 2008). For cooperative learning to be successful, students should be grouped carefully. In these groups, there should be students from different academic levels and different ethnicities with different learning habits (Maloof, & White, 2005).

One study stated that the topic "states of matter", one of the basic and most important subjects of a general chemistry course, could not be learned completely at macro and symbolic level, and that a complete learning of this subject was possible through micro-level learning rather than through macro and symbolic level learning (Yeziński, & Birk, 2006). To promote change in students' comprehension of chemistry topics at the micro level, researchers have implemented such teaching pedagogies as inquiry-based learning, cooperative learning, discrepant events, (social) constructivism, analogies, concrete models and visual tools/multiple representations. In recent years, using visual tools such as static or computer animated molecular models accompanied with oral and written discourse has gained prominence and was acknowledged to be promising in the construction of scientific conceptions (Adadan, 2006).

The purpose of this study was to investigate the effects of computer animations and cooperative learning on students' comprehension of chemistry topics (states of matter) at macro, micro, and symbolic levels. Specifically, the effects of animation (dynamic, two- and three-dimensional graphic representation) and cooperative learning on the comprehension of states of matter are examined at macro, micro, and symbolic levels. Differences in effects upon students of varying reasoning abilities are explored. The specific research question is:

Does instruction using traditional teaching, cooperative learning and computer animation create differences in students' comprehension of multiple meanings of the states of matter (macro and symbolic and micro levels)?

Method

Research Design

The experimental design for this study is a posttest-only control group designed for multiple meaning. This experimental design was chosen to eliminate any interaction that could occur between a pretest and the treatment for achievement. Although elimination of the pretest makes it impossible to show statistically that the groups are equal with respect to prior knowledge, it was important to avoid the experimental contamination that could have occurred from interaction between the pretest and the treatment. However, in the absence of a pretest, randomization should provide sufficient assurance that the groups lacked initial bias (Creswell, 2003).

Sample

The subjects of the present study comprised 64 university students enrolled in three classes of the general chemistry course taught in the first year of the elementary education science teacher education program during the 2006–2007 academic year. One class was randomly selected as the cooperative group (n=23: 9 female, 14 male), in which the cooperative learning method was applied. A second class was randomly selected as the computer animation group (n= 21: 8 female, 13 male), in which the computer animation technique was applied. The remaining class served as the control group (n=20: 8 female, 12 male), in which traditional teaching methods were applied. Pre-service science teachers are admitted to this department only after they have successfully passed a university entrance exam. The mean age of the participants was 19.47 (SD=1.44). Neither age nor gender differed significantly among the groups. Ages ranged from 18 to 23 years. Volunteers were given background information regarding the study prior to consent. During the training period, instruction for all groups was delivered by the researchers.

Instruments

Three dependent variables were measured: micro level meaning, non-micro level (macro and symbolic) meaning and students of varying reasoning abilities. Understanding at the micro level was determined by a States of Matter Test (SMT). Understanding at the non-micro level (macro and symbolic) was determined by a Course Exam (CE). Students' varying reasoning abilities were determined by a Test of Logical Thinking (TOLT).

The SMT is an instrument that requires students to make drawings, give explanations, and select the correct responses to multiple-choice questions. The categories of responses for the SMT were established by a panel of experts. Responses given in terms of molecules, atoms, ions and so on are classified as showing micro level understanding. Responses that repeat the question or are irrelevant or unclear are classified as showing no understanding. These responses were not taken into consideration for this study. The criteria and scale used in this study were developed by adapting the scale used for misconceptions by Haidar and Abraham (1991). For statistical analysis, numeric scores of '1' were assigned to satisfactory micro level responses and '0' to all other categories of responses. This

test was developed by the author and three chemistry teachers. The validity of the test was checked by a professor and two other chemistry teachers. The SMT was divided into five modules. Each module consisted of three questions (two open-ended, one multiple choice). This test gave continuous scale scores ranging from 0 to 15. A panel expert established the content validity, while the percent agreement for multiple graders on papers randomly chosen established the inter-rater reliability. The percentage agreement of the SMT was established at 77% or higher.

The CE had fifteen multiple-choice questions on states of matter, with each question worth five points. The CE was developed by the author and two chemistry teachers, and piloted with undergraduates from two college chemistry classes. Item analyses were performed for each question and confusing or vague questions were rewritten before the test was used in the study. After item analysis, nothing was changed in the scale so as not to violate its content validity. The CE was administered to students who had seen the relevant unit before to determine its reliability; the reliability co-efficiency (Cronbach Alpha) was found to be 0.69. Also for the validity of the developed CE, opinions of chemistry lecturers and researchers on the subject were taken into consideration. Researchers pointed out that the gains of the CE related to the subject of the states of matter have been high towards measurement.

The TOLT was used to identify potential differences in the cognitive skills of students and to control for this extraneous effect if it were influential on their learning. The TOLT, developed by Tobin and Capie (1981), is used to determine the formal reasoning ability of students. The test contains 10 items, eight two-tier multiple choice questions and two open-ended items. This test gives continuous scale scores ranging from 0 to 10. The internal reliability for this test is reported as 0.85 (Harwood, & McMahon, 1997). Tobin and Capie report a strong correlation of 0.80 between scores on the TOLT and formal reasoning skills, namely, controlling variables, and proportional, combinatorial, probabilistic, and correlation reasoning. The TOLT used for this study was translated into Turkish. The appropriateness of the questions to the Turkish language in terms of expression and meaning was analyzed by two instructors from the Turkish Language Teaching Department of Atatürk University Kazım Karabekir Education Faculty and the suggested corrections were made by the researcher. Consistency between the adapted and original versions was checked by two instructors at the Language Center of Atatürk University; these necessary revisions were made as well. The validity analysis, which aimed to find out whether the scale measured the targeted skills, was carried out by three experts from the departments of Elementary Mathematics Education and Elementary Sciences Education. Based on recommendations from these experts, two open-ended questions were extracted from the scale. Therefore, eight items of the original TOLT were used. This test gives continuous scale scores ranging from 0 to 8. The TOLT was administered to students who had seen the relevant unit before to determine its reliability. The reliability co-efficiency (Cronbach Alpha) was found to be 0.72. Examples of TOLT, SMT and CE questions are given in the appendix.

Animations

For students to understand the state of matter of a substance, it can be useful to use animations showing the intermolecular and intramolecular bonds that make up a chemical compound and the molecules of these compounds. For this reason, three main categories of animations were used in lectures during the teaching of states of matter. The animations used are summarized in Table 1.

Table 1

Animations Used in the State of Matter Instruction

Animation	Topic Shown	Length
Category 1	Molecular structure of chemical compounds	2 minutes
Category 2	Intermolecular and intramolecular bonds	2 minutes
Category 3	Matter phase and state change	2.5 minutes

Category 1 animations were prepared to demonstrate and teach the molecular structure of a chemical compound at the micro level. Category 2 animations were prepared to demonstrate and teach intermolecular and intramolecular bonds at the molecular level. Category 3 animations were prepared to demonstrate and teach states of matter. One example of a category 3 animation is given in Figure 1.

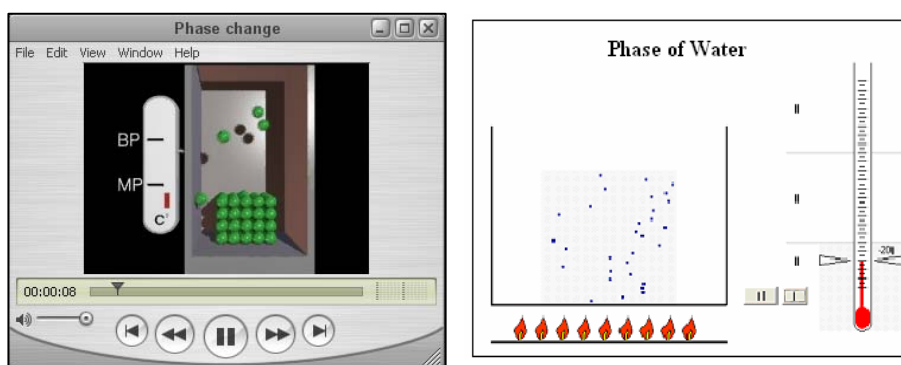


Figure 1. Snapshots of animations representing two different phase changes
http://mutuslab.cs.uwindsor.ca/schurko/animations/waterphases/status_water.htm

Procedure

The cooperative class was divided into five heterogeneous groups: two groups of four students and three groups of five students. Before the beginning of the instruction, the teacher gave information about learning objectives, the instruction process, rules for working in a cooperative group, group member roles, and assessment strategies (Johnson & Johnson, 1999). Students in the groups were encouraged to decide who would be the leader. Later, the heads of the groups were

determined by the group members. The subject of related states of matter was presented to the group members by the group heads. Each group studied their subject out of and in class. All activities were completed by students under the guidance of the teacher. While students were discussing in their small groups, the teacher visited all the groups and asked guiding questions to lead students in appropriate directions. All the cooperative groups prepared their own reports after the activities were completed. Each group was given 40 minutes to present their work in the classroom and 10 minutes for discussion with the class. During this discussion, the group answered questions from the class. All groups completed their topics in three weeks.

In the animation group, the researcher spent the first five minutes of the lesson asking questions of the class to determine their previous knowledge on the subject. Later, the subject was taught and the related animations were shown to the class over a period of 35 minutes. The animations were shown by projecting them onto a board, using a projection device compatible with computers. After the presentation of the animations, questions related to the subject were asked by the teacher for 10 minutes. Parts of the subject poorly understood were determined from the answers. These parts were covered again, using the animations.

In the control group, the subject was taught using traditional instructional methods. The researcher planned activities for the presentation of the subject that would be taught during the lesson. This was done not through classical teaching presentation style but by giving assignments to students on the subject of states of matter, and by providing workbooks for students to use to construct the information to be presented to them. The same content was taught in the animation and cooperative group and the learning objectives were the same. In contrast with the experimental groups, students in the control group were required to use their textbooks; students were passive participants and rarely asked questions; they did not benefit from the library or Internet sources; activities such as computer animations or brainstorming were not used. Instead, generally the teacher wrote concepts on the board and then explained them; students listened and took notes as the teacher lectured on the content. During this process, students' performances were observed and instruction was directed according to their feedback. The author taught states of matter to the treatment groups for three hours per week for three weeks. The TOLT, SMT, and CE were given to the treatment groups at the end of the study.

Data analysis

One-way analyses of covariance (ANCOVA) tests were used to analyze differences among the control, cooperative, and animation groups, with reasoning ability (TOLT) as a covariant. ANCOVA tests were formed for the CE and the SMT. Post-hoc tests were used to determine how the groups differed. Furthermore, descriptive statistics related to total mean scores of the TOLT, SMT, and CE were analyzed for the groups.

Findings and Results

Descriptive statistics related to the total mean scores of the TOLT, SMT, and CE for the treatment groups are presented in Table 2. According to this data, mean

scores of the groups range from 3.25 to 3.86, from 5.70 to 8.43, and from 36.50 to 38.81 for TOLT, SMT and CE, respectively.

Table 2
Descriptive Statistics for TOLT, SMT, and CE

GROUPS		TOLT	SMT	CE
Control	Mean	3.25	5.70	36.50
	N	20	20	20
	Sd. Deviation	1.86	1.95	8.13
Cooperative	Mean	3.70	7.74	36.96
	N	23	23	23
	Sd. Deviation	1.74	1.57	7.19
Animation	Mean	3.86	8.43	38.81
	N	21	21	21
	Sd. Deviation	1.53	0.98	8.79

To determine the relationships between the measures used, a correlation matrix of the TOLT scores, SMT scores, and CE scores was calculated for the states of matter topic. A lower correlation existed between the TOLT and SMT scores ($r = 0.06$). The correlation between the TOLT scores and the CE scores was $r = 0.34$. According to the data in Table 2, the mean TOLT scores of the groups varied, although not significantly.

One-way analyses of variance (ANOVA) tests were used to analyze differences among the control, cooperative, and animation groups based on reasoning ability. Data obtained from the TOLT are given in Table 3.

Table 3
ANOVA Results for TOLT Scores

	Sum of Squares	DF	Mean Square	F	p
Between Groups	4.043	2	2.022	0.688	0.506
Within Groups	179.191	61	2.938		
Total	183.234	63			

According to the data in Table 3, a one-way analysis of variance (ANOVA) found no significant differences in the TOLT scores of the control, animation, and cooperative groups [$F(2,63) = 0.688$; $p > 0.05$]. This finding supports the assumption that the groups should be considered equivalent. The TOLT results showed no differences among the treatment groups in this study (Table 3). This also helps to substantiate the assumption that the groups were similar. The TOLT mean scores in this study were higher than those found for college science students by Tobin and Capie (1981), but lower than those found for college chemistry students by Graves (1998). One explanation for this may lie in the fact that states of matter is generally

taught in the second half of the course. By this point, many students without the ability for proportional reasoning or controlling variables have dropped the course. These abilities are required for success in earlier units on stoichiometry, periodicity, chemical bonding, and atomic structure. Since attendance was required for inclusion in the study, it may be that a large proportion of students with low TOLT scores chose not to attend class and were not included in the study.

Micro level understanding was measured by the SMT. Each SMT was given during the week after the completion of the states of matter instruction. According to data in Table 2, the animation and cooperative groups' TOLT scores were 6.1% and 4.5% higher than that of the control group, respectively. The animation and cooperative groups' SMT scores were 37.3% and 20.4% higher than that of the control group, respectively. Therefore, the TOLT score was used as a covariate in the subsequent analyses to partial out its effects.

According to the data given in Table 4, the ANCOVA results show significant differences in the SMT scores by treatment even when the effects of the TOLT score are removed [$F(2, 63) = 16.793$; $p < 0.05$].

Table 4
ANCOVA Results for SMT

Source	DF	Mean Square	F	p
TOLT	1	0.142	0.058	0.810
Groups	2	40.761	16.793	0.000
Error	60	2.427		

Multivariate analysis (post-hoc testing) was used to determine where differences existed. The Games-Howell post-hoc test was chosen because it is robust to unequal cell sizes. The control group had SMT scores significantly different from those of the animation and the cooperative groups. The scores of the two experimental groups did not differ significantly. However, average mean scores of the animation group were higher than those of the cooperative group. Effect sizes calculated were 1.40 and 1.42 for the control group and the animation and cooperative groups, respectively. Effect sizes were calculated by dividing the difference in the means of the control group and one experimental group by the standard deviation of the control group.

Micro level understanding based on the SMT was related to the type of group (control, animation or cooperative) and to the TOLT scores (Table 4). One possible explanation for this might be that, for the simple micro level meanings depicted in this study, the maximum effect was achieved with both animation and cooperative learning. Another possible explanation may be that students only need to be cued to the dynamic particulate nature of these processes. The improved scores of the animation group are surprising when one considers that the animated sequences were short (2 to 2.5 minutes maximum). However, animations were used consistently for the duration of the instruction. Both of these facts add to the possibility that students, especially those with

high reasoning ability scores as in this study, may only need to be cued to internally visualize dynamic particle models.

The research on cooperative learning shows that the cooperative setting provides students with opportunities to engage in higher-order thinking skills and in processes of shared thinking, which helps them to not only gain a better understanding but also to build on their own contributions to develop new understandings and knowledge (Gillies, 2006; Hennessy & Evans, 2006; Lin, 2006). Students could not learn by only working in a small group. They need to construct their knowledge. Many students tend not to learn meaningfully, having difficulties relating what is taught to them with their real-world experiences and with other scientific ideas previously learned (Novak, 2002). Because of this, this study focused on the construction of knowledge in small cooperative groups.

Understanding at the non-micro level (macro and symbolic) was measured by the CE. The ANCOVA results for CE scores are given in Table 5.

Table 5
ANCOVA Results for CE

Source	DF	Mean Square	F	p
TOLT	1	32.248	0.582	0.449
Groups	2	9.375	0.169	0.845
Error	60	55.450		

According to these ANCOVA results, no significant differences were found between the groups in terms of CE scores [$F(2,63) = 0.169$; $p > .05$]. The CE showed no differences in understanding at the non-micro level (macro and symbolic) among the treatment groups (Table 5). Upon scrutiny of the questions on the CE, a possible explanation was found. The majority of the questions on the instructor-constructed exam were algorithmic in nature. The proposition that students memorized equations and the manipulation of equations that were needed to answer algorithmic problems without gaining micro level understanding may account for the lack of differences among the groups shown by the CE when very different results were found with the SMT scores.

Also, SMT responses indicate some interesting similarities and differences among the groups. The results (Q number 2 on STM) showed that the majority of the students (82, 76, and 72 percent of the control, cooperative and animation groups, respectively) did not understand the behavior of water molecules in ice at between -20 and 0 °C at the micro level. They did not consider that water can evaporate at every temperature. The main reason for this kind of learning difficulty could be due to textbooks' symbolic representations. This is interesting because, as Yore (1991) pointed out, science students usually see textbooks as very important sources of information that have a strong influence on shaping their teaching. Some responses given to this question by treatment group students are given in Figure 2.

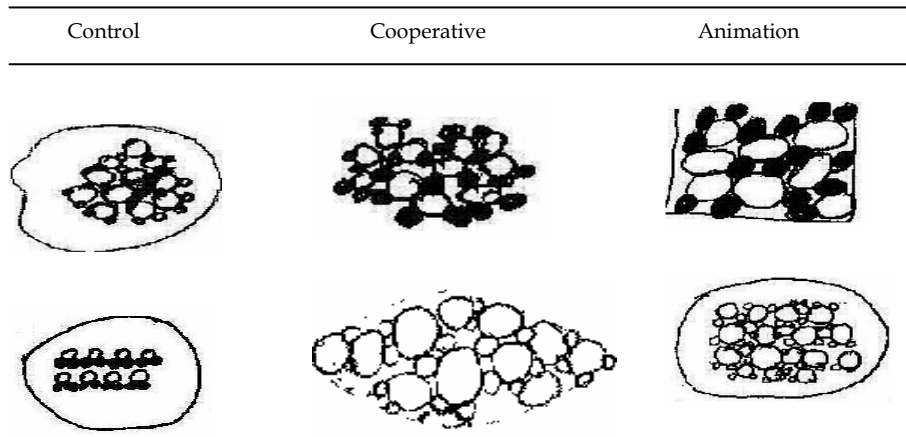


Figure 2. An example from treatment group students' responses concerning the behavior of water molecules in ice between -20 and 0 °C at the micro level

Other similarities in the SMT responses (Q number 4) included answers about the envisaging of molecules in the vapor phase of water; 75% of the control group, 82% of the cooperative group, and 88% of the animation group answered this question correctly. The main reason for this may be the fact that teachers give examples of the vaporization phase of water both in textbooks and on the Internet. Some responses given to this question by treatment group students are given in Figure 3.

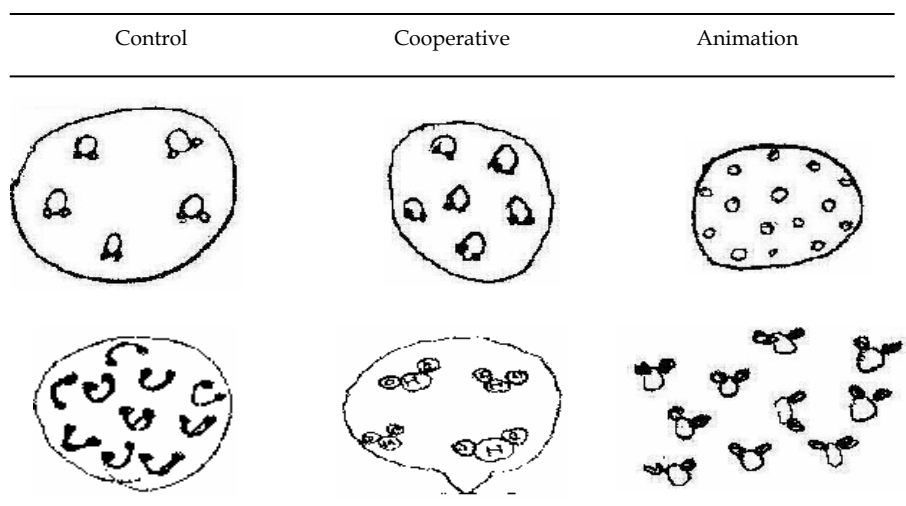


Figure 3. An example from treatment group students' responses concerning the molecules in the vapor phase of water

An interesting difference in the groups was found when the subjects were asked to draw a picture representing the change that occurs when a solid melts (Q number 3 on SMT). A few students chose to conserve particles between the SMT drawing of the solid and their drawing of the liquid. Only 54% of the control group conserved particles, while 65% of the cooperative group and 78% of the animation group conserved particles. Some treatment group students' responses to this question are given in Figure 4.

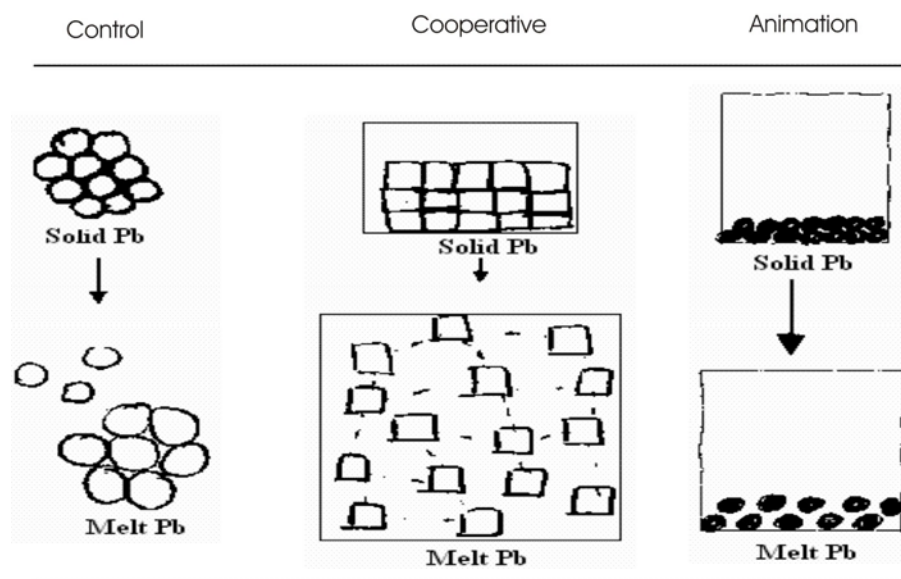


Figure 4. An example from treatment group students' responses concerning the molecules in solid lead and melted lead

Question number 5 on the SMT presented students with two sealed flasks: a blank flask and a flask filled with gas particles. The students were asked to draw the substance in the blank flask after it had been liquefied. Some students spontaneously depicted liquid-vapor equilibrium. 16% percent of the control group, 24% of the cooperative group, and 30% of the animation group drew the particles in a liquid phase with few particles in the gas phase. Some responses given to this question by treatment group students are given in Figure 5.

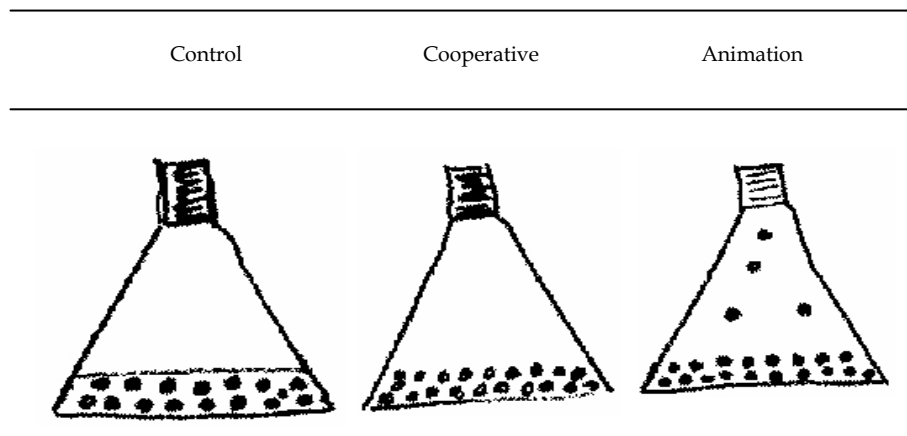


Figure 5. An example from treatment group students' responses concerning the view of water molecules in liquid phase

One possible explanation for these questions' (Q 3 and 5) findings is that use of animations caused the students to begin to think in particulate terms and to attend to more details concerning the behavior of particles. In the cooperative group, students gave more correct answers to these questions. This may be due to they took part in the learning process actively in both in-class and out-of-class discussions.

Conclusions and Recommendations

This study indicates that if cooperative learning instruction is organized, giving consideration to constructivism, students' achievement of micro level meanings will improve. The present paper could indicate a slight remediation for non-micro (macro and symbolic) level understanding. If the studies can become reality and if teachers can be encouraged to apply them in their classes, students' difficulties with micro meaning can be prevented. Thus, meaningful and effective learning can be provided.

In this study, animations provided a more scientifically correct visual model for submicroscopic processes not easily visualized. A few students viewing the animations had difficulties with the micro meaning as a consequence. Students who viewed the animations held a more particulate view of matter, ions dissolving in water, states of matter, and chemical bonding structure. More conservation of particles between drawings and fewer "continuous matter" drawings were evidence of this. The use of animations may increase understanding of micro meaning by prompting the formation of dynamic mental models of the phenomena. The dynamic quality of animation may promote deeper encoding of information than that of static pictures. Particle-level animations should be used frequently in chemistry classrooms to help students visualize particle-level behavior. In conjunction with showing these animations, students should be given opportunities to discuss and interpret the animations as they relate them to macroscopic phenomena that they have observed (Yezirski, & Birk, 2006).

Students from the cooperative learning group learned the unit related to the states of matter better than those who were in the group taught by the micro-level conventional method because students who studied on line with this method had the opportunity to utilize Internet facilities in addition to sources from the library. They were able to develop posters, pictures, simulations and animations of their own as an outcome. This was observed when the students in the cooperative learning group presented their research. Previous studies have reported that cooperative learning leads students to research using different sources and reconstruct their knowledge according to their own cognitive nature (Gillies, 2006; Hennessy, & Evans, 2006; Lin, 2006).

This study demonstrated that cooperative learning in small groups helped students understand the subject topic of the states of matter at the micro-level and develop cognitive models. In addition, conventional teaching methods based on teacher presentation were found to be as effective as cooperative learning and computer animation-aided instruction in students' ability to learn the states of matter at macro and symbolic levels. Researchers have shown that students experience more difficulty in micro-level understanding than they do in macro and symbolic levels (Wu et al., 2001). That students' micro-level understanding in the conventional teaching group was lower than that in the cooperative learning and animation groups might stem from the insufficiency of activities that could facilitate micro-level understanding. Consequently, when correct and suitable learning strategies are used, we think that it is more likely that the sources of macro level understanding, symbolic level understanding and misunderstandings will be remedied.

References

- Adadan, E. (2006). *Promoting high school students' conceptual understandings of the particulate nature of matter through multiple representations*, Unpublished Doctoral Dissertation, The Ohio State University.
- Barnea, N. & Dori, Y. J. (1996). Computerized molecular modeling as a tool to improve chemistry teaching, *Journal of Chemical Information and Computer Sciences*, 36 (6), 629-636.
- Bearison, D. J., Magzomes, S., & Filardo, E. K. (1986). Socio-cognitive conflict and cognitive growth in young children, *Merrill-Polmer Quarterly*, 32 (1), 51-72.
- Benson, D.L., Wittrock, M.C., & Bauer, M.E. (1993). Students' preconceptions of the nature of gases, *Journal of Research in Science Teaching*, 30 (6), 587-597.
- Ben-Zvi, R., Eylon, B., & Silberstein, J. (1988). Theories, principles and laws, *Education in Chemistry*, 25 (1), 89-92.
- Bowen, C. W. (2000). A Quantitative Literature Review of Cooperative Learning Effects on High School and College Chemistry Achievement, *Journal of Chemical Education*, 77 (1), 116-119.

- Creswell, J. W. (2003). *Research design qualitative, quantitative, and mixed method approaches* (2nd ed.), Thousand Oaks, CA: Sage.
- Doymuş, K. (2008). Teaching chemical equilibrium with the jigsaw technique. *Research in Science Education*, 26 (1), 47-57.
- Ebenezer, J. V. (2001). A hypermedia environment to explore and negotiate students' conceptions: Animation of the solution process of table salt, *Journal of Science Education and Technology*, 10 (1), 73-92.
- Eilks, I. (2005). Experiences and reflections about teaching atomic structure in a jigsaw classroom in lower secondary school chemistry lessons, *Journal of Chemical Education*, 82 (2), 313-319.
- Fleming, S. A., Hart, G. R., & Savage, P. B. (2000). Molecular orbital animations for organic chemistry, *Journal of Chemical Education*, 77 (6), 790-793.
- Friedel, A.W. & Maloney, D.P. (1992). An exploratory, classroom-based investigation of students' difficulties with subscripts in chemical formulas, *Science Education*, 76 (1), 65-78.
- Gabel, D. L. (1993). Use of the particle nature of matter in developing conceptual understanding. *Journal of Chemical Education*, 70 (3), 193-194.
- Gilbert, J. K. (2005). *Visualization in science education*, Netherlands: Springer.
- Gillies, R. M. (2006). Teachers' and students' verbal behaviors during cooperative and small-group learning, *British Journal of Educational Psychology*, 76 (2), 271-287.
- Graves, A. P. (1998). *An investigation comparing traditional recitation instruction to computer tutorials which combine 3-D animation with varying levels of visual complexity, including digital video in teaching various chemistry topics*, Unpublished Doctoral Dissertation, The University of Oklahoma Graduate College, Norman, Oklahoma.
- Greca, I.M. & Moreira, M.A. (2000). Mental models, conceptual models, and modeling, *International Journal of Science Education*, 22 (1), 1-11.
- Haidar, A.H. & Abraham, M.R. (1991). A comparison of applied and theoretical knowledge of concepts based on the particulate nature of matter, *Journal of Research in Science Teaching*, 28 (8), 919-938.
- Harwood, W. S. & McMahon, M. M. (1997). Effects of integrated video media on student achievement and attitudes in high school chemistry, *Journal of Research in Science Teaching*, 34 (8), 617-631.
- Hennessy, D. & Evans, R. (2006). Small-group learning in the community college classroom, *The Community College Enterprise*, 12 (1), 93-109.
- Johnson, D.W. & Johnson, R.T. (1999). Making cooperative learning work, *Theory Into Practice*, 38 (2), 67-73.
- Johnstone, A.H. (1991). Why is science difficult to learn? Things are seldom what they seem, *Journal of Computer Assisted Instruction*, 7 (1), 75-83.
- Johnstone, A.H. (1993). The development of chemistry teaching: A changing response to changing demand, *Journal of Chemical Education*, 70 (9), 701-705.
- Lin, E. (2006). Learning in the science classroom, *The Science Teacher*, 73 (1), 35-39.

- Maloof, J. & White, K. B. V. (2005). Team study training in the college biology laboratory, *Journal of Biological Education*, 39 (3), 120-124.
- Novak, J. D. (2002). Meaningful learning. The essential factor for conceptual change in limited or inappropriate prepositional hierarchies leading to empowerment of learners, *Science Education*, 86 (4), 548-571.
- Sanger, M. J., Phelps, A. J., & Fienhold, J. (2000). Using a computer animation to improve students' conceptual understanding of a can-crushing demonstration, *Journal of Chemical Education*, 77 (11), 1517-1520.
- Sanger, M.J. & Badger, S.M., (2001). Using computer-based visualization strategies to improve students' understanding of molecular polarity and miscibility, *Journal of Chemical Education*, 78 (10), 1412-1416.
- Shubbar, K. E. (1990). Learning the visualization of rotations in diagrams of three dimensional structures,, *Research in Science & Technological Education*, 8 (2), 145-155.
- Siegel, C. (2005). Implementing a research-based model of cooperative learning, *The Journal of Educational Research*, 98 (6), 339-351.
- Theall, R. M. (2003). *The effectiveness of computer-generated 3D animations in inquiry chemistry laboratory*, Unpublished Doctoral Dissertation, Arizona State University.
- Tobin, K. & Capie, W. (1981). The development and validation of a group test of logical thinking, *Educational and Psychological Measurement*, 41 (2), 413-423.
- Treagust, D.F., Chittleborough, G., & Mamiala, T.L. (2003). The role of submicroscopic and symbolic representations in chemical explanations, *International Journal of Science Education*, 25 (11), 1353-1368.
- Williamson, V. M. & Abraham, M. R. (1995). The effects of computer animation on particulate mental models of college chemistry students, *Journal of Research in Science Teaching*, 32 (5), 521-534.
- Wu, H.K., Krajcik, J.S., & Soloway, E. (2001). Promoting understanding of chemical representations: students' use of a visualization tool in the classroom, *Journal of Research in Science Teaching*, 38 (7), 821-842.
- Yair, Y., Mintz, R., & Litvak, S. (2001). 3D-virtual reality in science education: An implication for astronomy teaching, *Journal of Educational Multimedia and Hypermedia*, 20 (3), 293-305.
- Yeziarski, E. J. & Birk, J. P. (2006). Misconceptions about the particulate nature of matter using animations to close the gender gap, *Journal of Chemical Education*, 83(6), 954-960.
- Yore, L.D. (1991). Secondary science teachers' attitudes toward and beliefs about science reading and science textbooks, *Journal of Research in Science Teaching*, 28 (1), 55-72.

APPENDIX

State of Matter Evaluation Test (SMT)Question 6

Water exists in three forms (phases): gas (water vapor), liquid, and solid (ice).

Use circles to represent water particles in the boxes below.

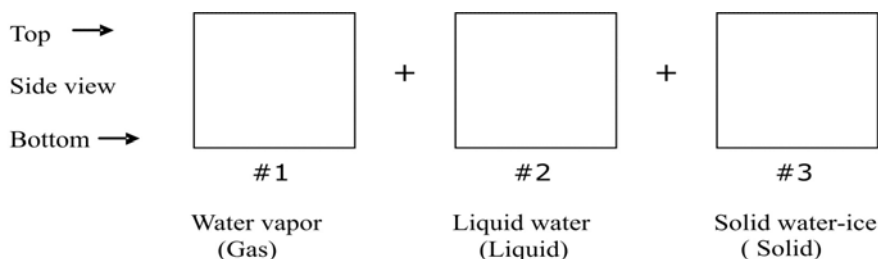


= water particle

In Box #1, draw 5 water particles in the gas phase.

In Box #2, draw 5 water particles in the liquid phase.

In Box #3, draw 5 water particles in the solid phase.

**Test of Logical Thinking (TOLT)**

For example, proportional reasoning is measured by the following pair of questions taken from the TOLT.

The Vegetable Seeds

Item 5

A gardener bought a package containing 3 squash seeds and 3 bean seeds. If just one seed is selected from the package what are the chances that it is a bean seed?

- a. 1 out of 2 b. 1 out of 3 c. 1 out of 4 d. 1 out of 6 e. 4 out of 6

Reason

1. Four selections are needed because the three squash seeds could have been chosen in a row.

2. There are six seeds from which one bean seed must be chosen.

3. One bean seed needs to be selected from a total of three.

4. One half of the seeds are bean seeds.

5. In addition to a bean seed, three squash seeds could be selected from a total of six.

Course Exam (CE)Item 4

Under the same conditions of temperature and pressure, a liquid differs from a gas because the particles of the liquid

1. are in constant straight-line motion
2. take the shape of the container they occupy
3. have no regular arrangement
4. have stronger forces of attraction between them
5. have more weak forces of attraction between them

Maddenin Hallerinin Mikro, Makro ve Sembolik Seviyede Öğrenilmesine Bilgisayar Animasyonları ve İşbirlikli Öğrenme Metotlarının Etkileri

(Özet)

Problem Durumu: Kimya eğitimi alan üniversite birinci sınıf öğrencilerinin çoğu kimya konularını makroskobik ve sembolik seviyelere nazaran mikroskobik seviyede anlamada zorluk çekmektedirler. Bu üç anlama seviyesi; 1) madde enerji ve doğa olayları gibi olguların anlaşıldığı makroskobik seviye, 2) molekül, atom, iyonlar ve buna benzer olguların anlaşıldığı mikroskobik seviye ve 3) formüller, eşitlikler ve iyon hareketleri gibi olguların anlaşıldığı sembolik seviye olarak ifade edilmektedir. Önemli olan bir husus bu üç seviyede öğrenmenin nasıl gerçekleştirileceği problemidir. Bu çalışmada, görsel modellerin kullanıldığı bilgisayar animasyonları ve öğrenci merkezli yöntemlerinden biri olan işbirlikli öğrenme metodunun öğrencilerin üniversite kimya eğitiminde maddenin halleri konusunu makro, mikro ve sembolik seviyede anlamalarına nasıl bir etki yapacağı araştırılmıştır.

Araştırmanın Amacı: Bu çalışmanın amacı, üniversite birinci sınıf öğrencilerinin maddenin halleri konusunu mikro, makro ve sembolik seviyelerde anlamaları üzerine bilgisayar animasyonları ve işbirlikli öğrenme metotlarının etkilerini belirlemektir.

Araştırmanın Yöntemi: Bu araştırmanın örneklemini 2006-2007 akademik yılında genel kimya dersini alan üç farklı sınıftaki toplam 64 fen bilgisi öğretmenliği birinci sınıf öğrencisi oluşturmuştur. Bu sınıflarda biri işbirlikli öğrenme metodunun uygulandığı işbirlikli grup, ikincisi bilgisayar animasyonları tekniğinin uygulandığı animasyon grubu ve üçüncüsü geleneksel öğretimin uygulandığı kontrol grubu olarak belirlenmiştir. Animasyon grubuna gösterilmek üzere ünite konularıyla ilgili hazırlanan bilgisayar animasyonları, dersin işleniş basamağında her bir animasyon iki dakikalık zaman içerecek şekilde sunulmuştur. Animasyon gösteriminin ardından konu ile ilgili sınıf tartışmaları yapılmış öğrencilerin yanlış ve eksik anlamaları olduğu durumlarda animasyonlar tekrar gösterilerek bu eksiklikleri giderici çalışmalar yapılmıştır. İşbirlikli öğrenme yönteminin uygulandığı sınıf 4-5 öğrenciden oluşan 5 heterojen gruba ayrıldı. Her gruba ilgili ünitenin konuları dağıtıldı. Gruplar konularını sınıf içerisinde ve sınıf dışarısında yapmış oldukları çalışmalarla hazırladılar. Hazırlıklar tamamlandıktan sonra her grup sınıf içerisinde 35 dakikalık sunum ve 15 dakikalık tartışma şeklinde sunumlarını tamamladılar. Kontrol grubunda ise ünite ile ilgili hazırlanan ders materyali doğrultusunda öğretmen sunumu soru-cevap vb. tekniklerin kullanıldığı geleneksel öğretim yöntemine göre ders işlendi. Çalışma bütün gruplarda üç haftalık süre zarfında yürütülmüştür. Araştırmada veri toplama aracı olarak Maddenin Halleri Testi (MHT), Ders Testi (DT) ve Mantıksal Düşünme Testi (MDT) kullanılmıştır. MDT, uygulamaya katılan öğrencilerin mantıksal düşünme düzeylerini tespit etmek amacıyla uygulanmıştır. Mantıksal düşünmenin dört alt boyutu ifade edilmektedir. MDT bu alt boyutları içine alan sekiz sorudan oluşmuştur ve değerlendirilirken doğru cevaplar 1 ve yanlış cevaplar 0(sıfır) puan olarak alınmıştır.

MHT, öğrencilerin maddenin halleri, bileşikler ve bağ yapılarına ilişkin soruları mikro seviyede çizimler yaparak göstermelerini içeren bir testtir. Bu testin puanlaması; mikro seviyede anlama puanları olarak değerlendirilmiştir. Puanlamada mikro seviyede memnun edici anlama puanlarına 1 puan verilmiş ve diğer cevap türleri puanlamaya alınmamıştır. Mikro seviye deki puanlar; moleküller, atomlar, iyonlar gibi terimlere verilen cevaplarından elde edilmiştir. DT, öğrencilerin maddenin halleri konusundaki makro ve sembolik (mikro seviyede olmayan) seviyedeki anlamalarını tespit etmek için kullanılmıştır. Bu test çoktan seçmeli 15 sorudan oluşturulmuştur. Testin güvenilirliğini belirlemek için daha önce maddenin halleri ünitesini görmüş olan öğrenci grubuna test uygulanmış ve güvenilirlik katsayısı 0,69 olarak bulunmuştur. Test sonuçları değerlendirilirken her bir sorunun doğru cevabı 5 puan olarak değerlendirilmiştir. Araştırma gruplarının DT ve MHT'den elde edilen puanları için tek yönlü kovaryans (ANCOVA) analizi kullanılmıştır.

Araştırmanın Bulguları

Mantıksal düşünme testinden elde edilen puanlara ait tek yönlü varyans (ANOVA) analizi sonuçları, kontrol, animasyon ve işbirlikli grupların mantıksal düşünme puanları arasında istatistiksel olarak anlamlı bir farklılık olmadığını göstermiştir. Bulgular araştırmaya katılan öğrencilerin mantıksal düşünme bakımından benzer özelliklere sahip oldukları varsayımını desteklemektedir. Maddenin halleri testinden elde edilen puanlara ait tek yönlü kovaryans (ANCOVA) analizi sonuçlarından, animasyon ve işbirlikli gruplardaki öğrencilerin kontrol grubuna göre daha yüksek mikro seviyede anlama puanlarına sahip oldukları bulunmuştur. Bununla birlikte maddenin halleri konusundaki mikro seviyede olmayan (makro ve sembolik seviyede) anlamalar bakımından araştırma grupları arasında anlamlı bir farklılık olmadığı tespit edilmiştir. Elde edilen bulgular hem bilgisayar animasyonlarının hem de işbirlikli öğrenmenin mikro seviyede anlamalar üzerinde önemli etkisinin olduğunu göstermiştir.

Araştırmanın Sonuçları ve Öneriler: Bu araştırma mikro seviyede gösterimler içeren bilgisayar animasyonları ile öğretimin öğretmen sunumuna dayalı geleneksel öğretime göre öğrencilerin kimya konularını mikro seviyede anlamalarını sağlamada daha etkili olduğu sonucuna varılmıştır. Bununla birlikte bu çalışma, küçük grup çalışmalarına dayalı işbirlikli öğrenme yönteminin de öğrencilerin maddenin halleri konusunu mikro seviyede anlamalarını ve zihinsel modeller oluşturmalarını sağladığını ortaya koymuştur. Ayrıca öğretmen sunumuna dayalı geleneksel öğretim yönteminin öğrencilerin maddenin halleri konusunu makro ve sembolik seviyede anlamalarını sağlamada işbirlikli öğrenme ve bilgisayar animasyonları ile öğretim kadar etkili olduğu sonucuna varılmıştır. Öğretmenler işbirlikli öğrenme aktivitelerini destekler ve kullanır iseler, öğrencilerin mikro seviyede anlama güçlüklerini önleyebilirler. Bu sayede etkili ve anlamlı öğrenme sağlanabilir. Sonuç olarak bizler doğru ve uygun öğrenme stratejisi kullanıldığında mikro seviyedeki anlamaların sağlanabileceğini düşünüyoruz.

Anahtar Sözcükler: Bilgisayar animasyonları, işbirlikli öğrenme, mikro seviye, maddenin halleri

Examining the Relationship between Epistemic Curiosity and Achievement Goals

Altay Eren*

Suggested Citation

Eren, A. (2009). Examining the relationship between epistemic curiosity and achievement goals. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 129-144.

Abstract

Problem statement: Although the prominent role of epistemic curiosity in students' achievement goals has long been emphasized in educational research, there is only one study in which the relationship between epistemic curiosity and achievement goals were empirically investigated. However, given the well-evidenced educational importance of both achievement goals and epistemic curiosity, it can be said that it is important to examine the relationship between them in order to explore the epistemic roots of students' achievement goals in settings, such as universities, which are achievement based. Thus, the present study has examined the relationship between epistemic curiosity and achievement goals.

Purpose of study: The aim of this study is twofold. First, it is to examine the relationship between students' achievement goals and epistemic curiosity; and second, it is to examine the role of epistemic curiosity as the predictor of students' achievement goals.

Method: Based on the survey method, a total of 309 undergraduate students voluntarily participated in the study. The Epistemic Curiosity Scale was used to assess students' curiosity as feeling of interest whereas the Curiosity as Feeling of Deprivation Scale was used to assess students' feelings of uncertainty and tension that motivated students' information-seeking and problem-solving behavior. Students' achievement goals, however, were assessed by the Achievement Goal Questionnaire. One-sample t tests, correlation, and hierarchical regression analyses were conducted to answer the research questions.

Findings and Results: The descriptive results of the study showed that both students' curiosity as a feeling of deprivation and as a feeling of interest were at relatively high levels. Students also tended to adopt both a performance-approach and a mastery goal over performance-avoidance goals. Contrary to the expectations, the results of the hierarchical

* Asst. Prof. Dr., Abant İzzet Baysal University Faculty of Education, eren_a@ibu.edu.tr

regression analysis demonstrated that not curiosity as a feeling of interest, but curiosity as a feeling of deprivation significantly predicted students' achievement goal orientations.

Conclusions and Recommendations: The present study revealed that students' epistemic curiosity, especially in the curiosity as a feeling of deprivation form, significantly correlated with students' mastery goals and performance-approach goals. Therefore, it is reasonable to suggest that the epistemic nature of students' curiosity, especially in the curiosity as a feeling of deprivation form, should be considered both in educational settings and future research.

Keywords: Epistemic curiosity, achievement goals, undergraduate students, hierarchical regression analysis

Achievement goals (AGs) are defined "as the competence-relevant purposes or aims that individuals strive in achievement settings" (Elliot, Shell, Henry, & Maier, 2005, p. 630). In other words, "achievement goals reflect desire to develop, attain or demonstrate competence in an activity" (Okun, Fairholme, Karoly, Ruehlman, & Newton, 2006, p. 255). Epistemic curiosity (EC), on the other hand, "reflects a desire for new information that motivates knowledge acquisition/exploratory behavior" (Litman, Hutchins, & Russon, 2005, p. 559), or shortly, it reflects "a drive to know" (Berlyne, 1954, p.187).

Despite the fact that both of these research areas are important in educational settings, few studies have directly examined the relationship between students' AGs and EC (Litman, 2008). More specifically, although the prominent role of EC in students' AGs was emphasized in previous research (Elliot, 1999), it has not directly been addressed to date (see Litman, 2008 for one exception). Given the well-evidenced educational importance of both AGs (e.g. Elliot, 1999) and EC (e.g. Chak, 2007), it can be said that it makes sense to examine the relationship between students' AGs and EC for at least one important reason. Accordingly, examining the EC as the predictor of students' AGs may help us in searching the epistemic roots of AGs in achievement-related settings such as universities.

Therefore, this study attempts to explore the relationship between students' AGs and EC. Specifically, the aim of this study is twofold: First, it is to examine the relationship between students' AGs and EC; and second, it is to examine the role of EC as the predictor of students' AGs. Based on this aim, the following research questions were formulated: (a) what are the AGs and EC held by the students? (b) Is there any significant relationship between students' AGs and EC? And (c) does students' EC significantly predict their AGs? Relevant concepts are summarized below.

Achievement Goals

Achievement goal theorists have described two achievement goals in particular: the goal to develop ability (variously labeled a task goal, learning goal, or mastery goal) and the goal to demonstrate ability or to avoid the demonstration of lack of ability (variously labeled an ability goal, ego goal, or performance goal) (Ames, 1992; Kaplan & Midgley, 1999; Midgley et al., 1998). However, Elliot and Church (1997)

proposed an integrative achievement goal conceptualization that includes both performance and mastery dimension as in the early form of theory, but a conventional performance goal was partitioned into an independent approach and avoidance components and three achievement goals were posited as mastery goals, performance approach-goals, and performance-avoidance goals.

Within this trichotomous framework, a mastery goal orientation is defined in terms of a focus on developing one's abilities, mastering a new skill, trying to accomplish something challenging, and trying to understand learning materials (Meece, Anderman, & Anderman, 2006). Performance-approach goals refer to an orientation towards demonstrating high ability whereas performance-avoidance goals refer to an orientation towards avoiding the demonstration of a low ability (Levy-Tossman, Kaplan, & Assor, 2007). More recently, Elliot and McGregor (2001) proposed an alternative framework comprising of four achievement goals. However, a trichotomous framework was adopted in the present study due to the reason that mastery goals, performance-approach goals, and performance-avoidance goals are the most evident goals in educational settings such as classrooms and schools (Régner, Escribe, & Dupeyrat, 2007).

A large body of research demonstrated that approach forms of AGs (i.e., mastery goals and performance-approach goals) are correlated with adaptive outcomes such as persistence and intrinsic motivation, course satisfaction, interest, and academic achievement, whereas the avoidance forms of AGs (i.e. performance-avoidance goals) are associated with maladaptive outcomes such as fear of failure, low interest, and poor academic performance (Dweck, 1986; Elliot & Church, 1997; Elliot & McGregor, 2001; Wolters, 2004; Ames & Archer, 1988; Shih, 2005; Gehlbach, 2006; Shen, Chen, & Guan, 2007). Furthermore, AGs can be adopted simultaneously in higher educational settings (i.e., colleges and/or universities) (Barron & Harackiewicz, 2001; Harackiewicz, Barron, Pintrich, Elliot, & Thrash, 2002). Barron and Harackiewicz (2001), for example, suggested that the adoption of both a performance-approach goal and a mastery goal can be more adaptive than the adoption of either a mastery goal or a performance-approach goal in the learning process.

Recently, based on the four-phase model of interest development (Hidi & Renninger, 2006) and the multiple goals model (Harackiewicz et al., 2002), Harackiewicz, Durik, Barron, Garcia, and Tauer (2008) examined the role of AGs in the development of interest, and found that initial interest, situational interest, class performance, and AGs were significantly related to one another. Despite the fact that Harackiewicz et al.'s research was not based on the EC framework, their research pointed out that the relationship between AGs and EC is not entirely speculative.

Epistemic Curiosity

The underlying cause of EC is one of the most basic problems that has occupied researchers. Loewenstein (1994, p. 93) proposed a theory through which he viewed "curiosity as occurring when an individual's informational reference point becomes elevated in a certain domain, drawing attention to an information gap" whereas

Spielberger and Starr (1994) conceptualized curiosity based on an optimal stimulation model through which it was suggested that optimal arousal is a function of both approaching new and unusual stimulation and avoiding unpleasant states of anxiety. Curiosity was defined as a feeling of deprivation (CFD) in the former whereas it was labeled as a feeling of interest (CFI) in the latter. The CFD reflects an “information gap” perspective, which can be defined by two quantities: “what one knows and what one wants or desires to know” (Loewenstein, 1994, p. 83).

According to Loewenstein (1994), such information gaps produce the feeling of deprivation-labeled curiosity. Within this view, the smaller the gap between these two quantities, the stronger one’s feeling of deprivation and thus, the higher the one’s desire to know. CFI, on the other hand, reflects an “interest” perspective, which involves positive feelings of interest associated with the anticipation of learning something new (Litman, 2005). These two perspectives on the nature of the EC are not incompatible (Litman & Jimerson, 2004). As Litman and Jimerson (2004, p. 148) explained, “Curiosity can reflect both the pleasurable anticipation of acquiring knowledge (i.e. CFI) and the feelings deprived by not having access to new information (i.e. CFD).” Therefore, EC was represented with both CFD and CFI in the present study.

Recently, Litman and Spielberger (2003) developed an Epistemic Curiosity Scale (EC-Scale), which has two factors—epistemic curiosity-diversive and epistemic curiosity-specific—which are used to assess individual differences in CFI. Epistemic curiosity-diversive assesses seeking a broad range of new information, whereas epistemic curiosity-specific measures interest in learning detailed knowledge about a specific topic. More importantly, they found that these two factors were captured by one higher-order factor, revealing that both factors of the scale can well be explained under the heading of CFI.

More recently, Litman and Jimerson (2004) developed a measurement of CFD (i.e. Curiosity as a Feeling of Deprivation Scale) (CFD-Scale) which has three factors: problem-solving, competence, and intolerance. Accordingly, the problem-solving subscale assesses problem-solving behavior motivated by feelings of tension or by a strong sense of urgency, whereas the competence subscale assesses the desire to reduce feelings of ignorance and a recognition of the value of feeling knowledgeable. The intolerance subscale, however, describes intolerance for situations in which information is inadequate. Similar to the EC-Scale, the subscales of the CFD-Scale were also captured by one higher order factor, suggesting that these factors can well be explained under the heading of CFD. Of particular interest, Litman and Jimerson also examined the relationship between CFD and CFI, and found that they are correlated significantly and positively with one another, signifying that EC could be conceptualized as involving feelings of both interest and deprivation.

Litman et al. (2005) investigated how knowledge-gaps, measured by the Feeling of Knowing, and individual differences in EC contribute to the arousal of state curiosity and exploratory behavior for undergraduate students. They found that the magnitude of perceived knowledge gaps measured by Feeling of Knowing and the EC trait both contributed to the arousal of curiosity states to motivate exploratory behavior when information was believed to be either unknown or at least partially

known. One of the interesting findings in Litman et al.'s study was that the participants wanted to check some of their answers regarding the items in general knowledge questionnaires even though they felt that they knew the answers. They concluded that such a motive may be labeled as curiosity, but it differs in its goal from the curiosity associated with a desire for intellectual enrichment, because if the answer was already known, there would be little new information to gain. According to Litman et al., this may be due to the relationship between participants' AGs and EC.

As a matter of fact, Litman (2005), in his research which, to this researcher's knowledge, is the only research on the topic, argues that CFI is highly related to scales that assess positive emotions such as mastery goals whereas CFD is somewhat more related to constructs that involve tension and anger such as performance-avoidance goals. In line with his argument, Litman (2008) recently found that CFI, as measured by the EC-Scale, was correlated with mastery-oriented learning whereas CFD, as measured by the CFD-Scale, was related to failure-avoidance and success orientation, pointing out that the students' approach and the avoidance gradient of achievement goals are rooted to EC in predicted direction.

In light of those explanations above, it was hypothesized that CFD will correlate with the avoidance form of AGs (i.e., performance-avoidance goals), whereas CFI will correlate with mastery goals. No specific hypothesis was suggested regarding the relationships among performance-approach goals, CFD, and CFI, due to the mixed findings about the adaptive role of performance-approach goals (Midgley, Kaplan, & Middleton, 2001).

Method

Participants

Based on the survey method, a total of 309 undergraduate students (159 females, 150 males), who majored in Business Administration ($n = 129$), Physics ($n = 82$), and Psychological Counseling and Guidance ($n = 98$) in a large university which is located in the northwest of the Black Sea Region in Turkey, voluntarily participated in the study. Students were recruited from the three domains which were randomly selected among the departments of the three large faculties of the university where the present study was carried out (i.e., faculty of education-approximately 4574 students, faculty of science and letters-approximately 1000 students, and faculty of management and business administration-approximately 2292 student). Of these 309 students, thirty-three were freshmen, ninety-seven were sophomores, eighty-four were juniors and ninety-five were seniors. Participants ranged in age from 17 to 33 years ($M = 21.21$, $SD = 1.83$). In Turkey, the cost of university education is relatively high for undergraduate students. Therefore, it was assumed that the level of the sample students' socio-economic status was at least medium.

Research Instruments

All items in the research instruments were translated into Turkish by the researcher with the assistance of three lecturers in the foreign languages department

of the university where the present study was carried out. Specifically, items were translated into Turkish by the researcher first, and then the lecturers translated them back into English. Finally, the original items were compared to those that were translated. Overall, the agreement rate on the items of the scales was quite high (88%). Disagreements were resolved through discussion of the items.

Epistemic Curiosity Scale. Because it has been widely used in curiosity research, the EC-Scale (Litman & Spielberger, 2003) was used to assess students' CFI in the present study. The EC-Scale is a 10-item scale that has two factors as epistemic curiosity-diversive (e.g., I enjoy learning about subjects which are unfamiliar) and epistemic curiosity-specific (e.g., I'm interested in discovering how things work). The EC-Scale measures CFI, namely, positive feelings of interest in gaining intellectual knowledge (Litman & Pezzo, 2007). As in the original scale, participants were asked to report how they generally felt on a 4-point Likert-type scale, ranging from 1 (almost never) to 4 (almost always). The internal reliability of the scale was .80 in the present study.

Curiosity as a Feeling of Deprivation Scale. The CFD-Scale (Litman & Jimerson, 2004) was used in the present study to assess students' feelings of uncertainty and tension that motivated students' information-seeking and problem-solving behavior. As mentioned earlier, CFD-Scale is a 15-item scale that has three factors in problem-solving (e.g., conceptual problems keep me awake thinking about solutions), competence (e.g., important to feel knowledgeable) and intolerance (e.g., critical of ideas and theories). Although the CFD-Scale and EC-Scale shared the same focus on desire to obtain intellectual knowledge, the former emphasizes acquiring knowledge in order to reduce feelings of tension due to uncertainty (Litman & Pezzo, 2007). The CFD-Scale also has a 4-point response format, ranging from 1 (almost never) to 4 (almost always). Cronbach's coefficient alpha was .85 in the present study.

Despite the fact that the CFD-Scale has three factors and EC-Scale has two factors, these factors are yielded to higher-order factors such as CFD and EC respectively, suggesting that the scales can also be used based on their higher-order factors. Thus, regardless of their respective factors, the sum of the item scores in both the CFD-Scale and the EC-Scale were considered in order to focus on their core aspects (i.e., feelings of interest and deprivation). As a result, the confirmatory factor analysis¹ showed that both one-factor CFD-Scale with 15 indicators (GFI = .95; AGFI = .94; CFI = .95) and one-factor EC-Scale with 10 indicators (GFI = .94; AGFI = .91; CFI = .91) has an acceptable fit to the data (Bollen & Curran, 2006).

Achievement Goal Questionnaire. The Achievement Goal Questionnaire (AG-Questionnaire) (Elliot & McGregor, 2001) was used to assess students' AGs in this study. The scale is comprised of mastery-approach, mastery-avoidance performance-approach, and performance-avoidance goal dimensions, each of which has three items. However, in accordance with the trichotomous framework, mastery-approach (hereafter mastery goals), performance-approach, and performance-avoidance goal dimensions of the scale were considered in the present study.

¹ Because the data were somewhat skewed, the Maximum Likelihood Estimation method (MLE) was not used; the Ordinary Least Squares method (OLS) from STATISTICA 7 was used in all the CFA analyses (Bollen & Curran, 2006).

As in the original scale, items were presented in a 7-point Likert scale format, ranging from 1 (not at all true of me) to 7 (very true of me). The internal reliabilities were .93, .84, and .77 for performance-approach goals, mastery goals, and performance-avoidance goals respectively.

However, it was observed that the reliability coefficient for the performance-avoidance scale would have improved to .89 if one of the items had been deleted (e.g., my fear of performing poorly in this class is often what motivates me). Therefore, performance-avoidance goal dimension was represented with two items. Finally, the confirmatory factor analysis showed that the three-factor model had a good fit to data (GFI = .99; AGFI = .98; CFI = .99).

Procedure

Data were obtained during the fall semester of the 2007-2008 academic year. The scales were applied with an interval of a few days in order to prevent a possible response bias. Students were asked to write a pseudonym on each scale in order to make sure that the scales are applied to each participant in the sample. Based on a match/mismatch procedure, only those students whose pseudonyms consistently appear on each of the scales were included in the sample of the study (n = 309) whereas the remaining students were excluded (n = 18). Furthermore, these scales were presented to participants with instructions concerning both the aim of the study and a brief explanation about the constructs. Gender, age, and year of study were assessed by a self-report on each of the scales. Administration lasted approximately 10-15 minutes for the EC-Scale, 15-20 minutes for the CFD-Scale, and 10-15 minutes for the AG-Questionnaire.

Data Analyses

Data were analyzed by SPSS 15.0 and STATISTICA 7 software. As preliminary analysis, the interactions among background variables (i.e., gender, age, year of study, and fields of study), CFD, and CFI were checked due to the possibility that the interactional effects of these variables on dependents may cause an artificial decrease or increase in the predictive power of independents. Therefore, a Multivariate Analysis of Covariance (MANCOVA) was conducted (Stevens, 1996). In the MANCOVA, AGs were entered as dependents and age as a covariate whereas gender, year of study, and fields of study were entered as independents. Following the MANCOVA, a series of Univariate Analysis of Variance (ANOVA) was also conducted. However, no significant interactions among the background variables, CFD, and CFI were detected. It can be said that, at least in the present sample, the possible effects of CFD and CFI on dependents would not be artificial and would not originate from the hidden effects of those relationships among background variables, CFD, and CFI. Thus, the interactional effects were not included in the analysis. Nevertheless, background variables were included in the analysis in order to control their direct effects on dependents. Indeed, these variables are frequently controlled in

educational research in order to obtain more robust and reliable results regarding the variables at hand (Stevens, 1996).

Utilizing a series of one-sample t-tests, students' CFD, CFI, and AGs were defined. Midpoints of the rating scales were 4 for Achievement Goal Questionnaire, and 2.5 for both curiosity scales. These midpoints were taken as reference points to compare the group means of each dimension of the scales. Addressing the second research question, both zero-order correlation (Pearson) and hierarchical regression analyses were conducted. Gender (females equal to 1 and males equal to 0), age (reported as open-ended), year of study, and fields of study variables were all entered in the first step of the regression analyses in order to control their possible effects on dependents (i.e. achievement goals). A traditional coding procedure was utilized for years of study and fields of study variables (Pedhazur, 1997). Accordingly, seniors were determined as an outer category for year of study, whereas Psychological Counseling and Guidance was the external category for fields of study. Thus, the remaining categories in fields of study and year of study were all compared to respective outer categories. Finally, CFI was entered in the second step of the analyses whereas CFD was entered in the final step of the analyses in order to see their main effects on dependents separately.

Results

Descriptive Analysis

The results of the one-sample t-tests were summarized in Table 1.

Table 1
Summary of the One-Sample t-Test Results

Variable	M (SD)	t	p
CFD	44.24 (6.91)	17.15	.000
CFI	28.69 (5.25)	12.35	.000
MAS	16.77 (3.18)	26.35	.000
PAP	13.94 (4.16)	8.17	.000
PAV	7.24 (3.30)	-4.05	.000

Note. CFD: Curiosity as a Feeling of Deprivation;
CFI: Curiosity as a Feeling of Interest; MAS:
Mastery Goals; PAP: Performance-Approach Goals;
PAV: Performance-Avoidance Goals.

As seen in Table 1, both means of CFD ($t(308) = 17.15, p < .001$) and CFI ($t(308) = 12.35, p < .001$) were significantly above the midpoint, indicating that both roots of the epistemic curiosity of sample students were at relatively high levels. Students' mastery goals ($t(308) = 26.35, p < .001$) and performance-approach goals ($t(308) = 8.17, p < .001$) were also significantly above the midpoint. However, the same was not true for students' performance-avoidance goals ($t(308) = -4.05, p < .001$).

Correlation Analysis

The results of the correlation analysis were presented in Table 2.

Table 2
Bivariate Correlations

Variable	1	2	3	4	5
1. CFD	-				
2. CFI	.44**	-			
3. MAS	.27**	.20**	-		
4. PAP	.13*	.09	.16*	-	
5. PAV	-.12*	-.10	-.28**	-.09	-

**p<.01; *p<.05

As shown in Table 2, CFD was positively and moderately correlated with mastery goals ($r = .27$, $p < .01$) whereas it was positively and weakly correlated with performance-approach goals ($r = .13$, $p < .05$). In addition, the relationship between the performance-avoidance goals and CFD was weak and negative ($r = -.12$, $p < .05$). CFI, on the other hand, correlated only with mastery goals ($r = .20$, $p < .01$). CFD and CFI strongly and positively correlated with one another ($r = .44$, $p < .01$). Mastery goals weakly and positively correlated with performance-approach goals ($r = .16$, $p < .05$), but moderately and negatively correlated with performance-avoidance goals ($r = -.28$, $p < .01$). Finally, no significant relationship was observed between the performance-approach and performance-avoidance goals ($r = -.09$, $p > .05$).

Hierarchical Regression Analysis

The summary of the regression results was presented in Table 3. Regardless of the effects of background variables (see Table 3), the second step of the analyses revealed that the CFI significantly predicted the students' mastery goals ($\beta = .19$, $p < .01$). However, in the third step, the effect of the CFI disappeared due to the CFD effects on mastery goals ($\beta = .23$, $p < .001$). Additionally, the CFD also predicted students' performance-approach goals ($\beta = .13$, $p < .05$). Finally, neither CFD nor CFI predicted students' performance-avoidance goals whereas the approach forms of AGs were significantly linked to students' CFD and CFI (see Table 3).

Discussion

The descriptive results of the study showed that both students' CFD and CFI were at relatively high levels. Given that the EC reflects a desire for new information that motivates knowledge acquisition/exploratory behavior (Litman et al., 2005), and such behaviors are usually reinforced in the university where the epistemic aspect of the curiosity is more salient, they explain why sample students' EC were at relatively high levels. The t-test results, on the other hand, revealed that both performance-approach

Table 3
The Summary of the Hierarchical Regression Results

Variable	MAS				PAP				PAV			
	B	S. E.	β	p	B	S. E.	β	p	B	S. E.	β	p
Step 1												
Age	-.05	.14	-.03	.732	-.30	.18	-.13	.086	.25	.14	.14	.068
Gender	.69	.37	.11	.063	.33	.47	.04	.481	.01	.37	.00	.986
Freshmen	.02	.91	.00	.979	-1.43	1.17	-.11	.222	1.83	.91	.17	.046*
Sophomores	-.32	.58	-.05	.587	-.92	.75	-.10	.221	1.22	.59	.17	.039*
Juniors	.48	.55	.07	.384	.33	.71	.04	.645	-.04	.55	-.01	.940
Business	.27	.46	.04	.559	2.44	.59	.29	.000***	1.77	.46	.27	.000***
Physics	1.38	.61	.19	.024*	1.84	.79	.20	.020*	-.89	.61	-.12	.147
Step 2												
Age	-.05	.13	-.03	.686	-.31	.18	-.13	.081	.25	.14	.14	.063
Gender	.74	.36	.12	.040*	.37	.47	.04	.436	-.02	.37	-.00	.948
Freshmen	-.14	.89	-.01	.875	-1.52	1.17	-.11	.192	1.91	.91	.18	.037*
Sophomores	-.33	.57	-.05	.571	-.92	.75	-.10	.218	1.22	.58	.17	.037*
Juniors	.31	.54	.04	.566	.23	.71	.02	.749	.05	.55	.01	.935
Business	.30	.45	.05	.506	2.46	.59	.29	.000***	1.75	.46	.26	.000***
Physics	1.33	.60	.19	.028*	1.81	.78	.19	.021*	-.87	.61	-.12	.159
CFI	.12	.03	.19	.001**	.07	.04	.09	.121	-.06	.04	-.10	.086
Step 3												
Age	-.10	.13	-.06	.443	-.34	.18	-.15	.051	.27	.14	.15	.048*
Gender	.47	.36	.07	.196	.16	.48	.02	.742	.09	.38	.01	.837
Freshmen	-.34	.88	-.03	.699	-1.68	1.16	-.13	.150	1.98	.91	.19	.030*
Sophomores	-.39	.56	-.06	.491	-.97	.74	-.11	.193	1.24	.58	.18	.034*
Juniors	.27	.53	.04	.611	.20	.70	.02	.782	.06	.55	.01	.913
Business	.45	.44	.07	.312	2.57	.58	.31	.000**	1.70	.46	.25	.000***
Physics	1.52	.59	.21	.011*	1.96	.78	.21	.013*	-.93	.61	-.13	.129
CFI	.05	.04	.09	.147	.02	.05	.03	.661	-.04	.04	-.06	.343
CFD	.10	.03	.23	.000***	.08	.04	.13	.038*	-.04	.03	-.08	.200

*** $p < .001$; ** $p < .01$; * $p < .05$

Note. MAS: $R^2 = .04$, $F(7, 301) = 1.84$, $p > .05$ for step 1; $R^2 = .08$, $F(8, 300) = 3.13$, $p < .01$ for step 2; $R^2 = .12$, $F(9, 299) = 4.34$, $p < .001$ for step 3. PAP: $R^2 = .08$, $F(7, 301) = 3.52$, $p < .01$ for step 1; $R^2 = .08$, $F(8, 300) = 3.40$, $p < .01$ for step 2; $R^2 = .10$, $F(9, 299) = 3.54$, $p < .001$ for step 3. PAV: $R^2 = .10$, $F(7, 301) = 4.89$, $p < .001$ for step 1; $R^2 = .11$, $F(8, 300) = 4.67$, $p < .001$ for step 2; $R^2 = .12$, $F(9, 299) = 4.34$, $p < .001$ for step 3.

goals and mastery goals were highly pursued by the students. However, the opposite was the case for students' performance-avoidance goals. This result is in accordance with the multiple-goal framework in which both performance-approach goals and mastery goals were adopted simultaneously (Barron & Harackiewicz, 2001). In fact, the results of the correlation analysis supported this claim. In other words, correlation analysis provided a significant framework in which CFD, as compared to CFI, was the dominant EC dimension that associated with students' AGs, suggesting that a multiple goal framework was the case for the sample students.

The results of the regression analysis, on the other hand, showed that it wasn't CFI, but CFD that significantly predicted students' AGs. More specifically, regardless of the effects of the background variables, CFD significantly predicted both students' mastery goals and performance-approach goals, indicating that students' mastery goals, as well as their performance-approach goals, were significantly affected by their CFD. This result is not in accordance with Litman's (2008) study. Given the fact that mastery goal-oriented students focus on learning for its own sake in achievement settings, one may intuitively expect that CFI, as a curiosity motive, is more reasonable or is at least more appropriate for mastery-goal-oriented students than CFD. However, according to Loewenstein (1994) "filling information gap is an important aspect of achieving competence, and curiosity is certainly and particularly strong when it comes to knowledge pertaining to one's own competence" (Loewenstein, 1994, p. 93). Loewenstein (1994) also suggested that CFD is a more powerful motive for learning than CFI. If this is the case, one can understand why students' mastery goals are CFD-driven. Mastery goal-oriented students are more interested in gaining knowledge and seeking information (Elliot, 1999). This makes them relatively well-informed about the tasks at hand. As demonstrated in the previous studies, the more a person approaches an answer, the more s/he wants to know the answer due to the reduction of tension attributed to uncertainty (Litman, 2005; 2008).

The same reasoning may be used to explain why CFD underlies the students' performance-approach goals because information sought during CFD reactions is theorized to be substantive, meaningful, and capable of increasing subjective feelings of competence such as solutions to a difficult problem (Litman, 2005), which is a desirable outcome for demonstrating the ability to others. Nevertheless, cultural background variables such as individualism and collectivism (e.g. Nisbett, 2005), which were not considered in the present study, may also have played a significant role in the emergence of different pictures regarding the links among Turkish undergraduate students' CFD, CFI, and AGs and those of their American peers. This issue deserves further investigation.

Following those explanations above, it can be said that both students' mastery goals and performance-approach goals were affected by CFD regardless of the focus of their AGs (e.g., performing better than the others or task mastery), indicating that not the focus, but the approach gradient is the case. As a matter of fact, CFD, like CFI, did not significantly predict students' performance-avoidance goals which lie on the avoidance gradient of AGs. In short, it seems that the mastery goal-oriented and performance-approach goal-oriented students' EC points more to CFD than to CFI. This result has broadened our current view about the curiosity motives underlying the students' mastery goals and performance-approach goals. Accordingly, CFD seems to be more important

than CFI in achievement goal-related settings such as university. Therefore, it can be said that considering CFD as a significant antecedent of students' mastery goals and performance-approach goals requires designing a learning environment in which students act as autonomous learners to fill the gaps between what they know and what they desire to know in order to take advantage of their mastery goals and performance-approach goals. Results also showed that CFD and CFI are strongly correlated with one another. This relationship may indicate a co-activation or interplay between them in achievement goal-related settings; this idea needs further investigation.

Limitations

This study has several limitations. First, the sample size was relatively small which limited the generalizability of the current results. Second, the data were correlational in nature, indicating that the results of this study could not be interpreted in a causative manner. Third, the participants' performance-avoidance goals were assessed only with two items. Although the performance-avoidance goal scale with two items had high internal reliability, this can also be seen as a limitation due to the small number of items in the scale. Finally, obtaining the data from only one university was another limitation of the study. Therefore, the results of this study should be cautiously interpreted in future research.

Conclusions and Recommendations

The present study revealed that (a) both students' CFD and CFI are at relatively high levels, suggesting that both dimensions of EC are highly represented in educational settings such as university; (b) students tend to adopt both performance-approach goals and mastery goals compared to performance-avoidance goals, signifying that it is more appropriate to focus on students' AGs within a multiple-goal framework, which is also replicated through the results of the correlation analysis; (c) the sample students' EC, especially in CFD form, is correlated with their mastery goals and performance-approach goals, pointing out that CFD is a significant antecedent of students' mastery goals and performance-approach goals.

Based on the results of this study, it can be suggested that (a) the epistemic nature of the students' curiosity should be considered both in educational settings such as university and in future research to broaden our current understanding with respect to the antecedents of AGs, as well as to broaden our viewpoint regarding the effects of learning environments on students' EC and AGs which have long been valued, but separately emphasized to date; and (b) not only the interest type of students' EC, but also the deprivation type, which seems more effective on students' achievement motivation, should be emphasized in order to activate and/or sustain curiosity in educational settings, which, in turn, may enhance students' mastery goals and performance-approach goals.

References

- Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261-271.
- Ames, C. & Archer, J. (1988). Achievement goals in the classroom: student learning strategies and achievement motivation. *Journal of Educational Psychology, 18*, 409-414.
- Barron, K. E. & Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: testing multiple goal models. *Journal of Personality and Social Psychology, 80*, 260-267.
- Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology, 45*, 180-191.
- Bollen, K. A. & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. New Jersey, NJ: Wiley-Interscience.
- Chak, A. (2007). Teachers' and parents' conceptions of children's curiosity and exploration. *International Journal of Early Years Education, 15* (2), 141-159.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040-1048.
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 169-189.
- Elliot, A. J. & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72* (1), 218-232.
- 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., Shell, M. M., Henry, K., & Maier, M. (2005). Achievement goals, performance contingencies, and performance attainment: An experimental test. *Journal of Educational Psychology, 97*, 630-640.
- Gehlbach, H. (2006). How changes in students' goal orientations relate to outcomes in social studies. *The Journal of Educational Research, 99* (6), 358-370.
- Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: necessary and illuminating. *Journal of Educational Psychology, 94* (3), 638-645.
- Harackiewicz, J. M., Durik, A. M., Barron, K. E., Garcia, L. L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology, 100* (1), 105-122.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 111*-127.
- Kaplan, A. & Midgley, C. (1999). The relationship between perceptions of the classroom goal structure and early adolescents' affect in school: the mediating role of coping strategies. *Learning and Individual Differences, 11* (2), 187-212.
- Levy-Tossman, I., Kaplan, A., & Assor, A. (2007). Academic goal orientations, multiple goal profiles, and friendship intimacy among early adolescents. *Contemporary Educational Psychology, 32* (2), 231-252.
- Litman, J. A. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition and Emotion, 19* (6), 793-814.

- Litman, J. A. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences, 44* (7), 1585-1595.
- Litman, J. A. & Jimerson, T. L. (2004). The measurement of curiosity as a feeling of deprivation. *Journal of Personality Assessment, 82* (2), 147-157.
- Litman, J. A. & Pezzo, M. V. (2007). Dimensionality of interpersonal curiosity. *Personality and Individual Differences, 43*, 1448-1459.
- Litman, J. A. & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diverse and specific components. *Journal of Personality Assessment, 80* (1), 75-86.
- Litman, J. A., Hutchins, T. L., & Russon, R. K. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behavior. *Cognition and Emotion, 19* (4), 559-582.
- Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116* (1), 75-98.
- Meece, J. L., Anderman, E. M., & Anderman, H. L. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57*, 487-503.
- Midgley, C., Kaplan, A., & Middleton, M. J. (2001). Performance-approach goals: good for what, for whom, under what circumstances, and at what cost? *Journal of Educational Psychology, 93*, 77-86.
- Midgley, C., Kaplan, A., Middleton, M., Maehr, M. L., Urdan, T., Anderman, L. H., Anderman, E., & Roeser, R. (1998). The development and validation of scales assessing students' achievement goal orientations. *Contemporary Educational Psychology, 23*, 113-131.
- Nisbett, R. E. (2005). *Düşüncenin coğrafyası: Doğulular ile Batılılar nasıl ve neden farklı düşünürler?* [The geography of thought: How Asians and Westerners think differently-and why? Tr. G. C. Güven]. İstanbul: Varlık.
- Okun, M. A., Fairholme, C., Karoly, P., Ruehlman, L. S., & Newton, C. (2006). Academic goals, goal process cognition, and exam performance among college students. *Learning and Individual Differences, 16*, 255-265.
- Pedhazur, E. J. (1997). *Multiple-regression in behavioral research: Explanation and prediction*. New York, NY: Wadsworth.
- Régner, I., Escribe, C., & Dupeyrat, C. (2007). Evidence of social comparison in mastery goals in natural academic settings. *Journal of Educational Psychology, 99* (3), 575-583.
- Shen, B., Chen, A., & Guan, J. (2007). Using achievement goals and interest to predict learning in physical education. *Journal of Experimental Education, 75* (2), 89-108.
- Shih, S. S. (2005). Role of achievement goals in children's learning in Taiwan. *Journal of Educational Research, 98* (5), 310-319.
- Spielberger, C. D. & Starr, L. M. (1994). Curiosity and exploratory behavior. In H. F. O'Neil, Jr. & M. Drillings (Eds.), *Motivation: Theory and Research* (pp. 221-243). Hillsdale, NJ: Lawrence Erlbaum.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Wolters, C. (2004). Advancing achievement goal theory: using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology, 96* (2), 236-250.

Bilgi Kaynaklı Merak ve Başarı Amaçları Arasındaki İlişkinin İncelenmesi

(Özet)

Problem Durumu: Konuyla ilgili literatürde, bilgi kaynaklı merak ve başarı amaçları arasındaki ilişkiye yönelik çıkarsamaların yapıldığı bazı araştırmalar bulunmasına rağmen, konunun doğrudan incelendiği yalnızca bir araştırmaya rastlanmıştır. Ancak, hem başarı amaçları hem de bilgi kaynaklı merak konusunun iyi bilinen önemi dikkate alındığında, bu değişkenler arasındaki ilişkiye yönelik bir çalışmanın üniversiteler gibi başarı-amaçlarının önemli bir etken olduğu eğitim/öğretim ortamlarında öğrencilerin başarı amaçlarının bilgi kaynaklı merakla ilgili kökenlerinin keşfedilmesi açısından önemli olduğu söylenebilir. Dolayısıyla, bu çalışma söz konusu iki değişken arasında bugüne kadar ihmal edilmiş boşluğu doldurabilmek için “öğrencilerin başarı amaçları ve bilgi kaynaklı merakları nedir?”, “Öğrencilerin başarı amaçları ve bilgi kaynaklı merakları arasında anlamlı bir ilişki var mıdır?” Ve “öğrencilerin bilgi kaynaklı merakları başarı amaçlarını anlamlı düzeyde yordamakta mıdır?” sorularına odaklanmıştır.

Araştırmanın Amacı: Bu araştırmanın amacı iki yönlüdür: Birincisi, öğrencilerin başarı amaçları ve bilgi kaynaklı merakları arasındaki ilişkinin incelenmesi, ikincisi ise bilgi kaynaklı merakın öğrencilerin başarı amaçlarındaki rolünün incelenmesidir.

Araştırmanın Yöntemi: Tarama yönteminden hareketle, Türkiye'nin Batı Karadeniz bölgesinde yer alan büyük bir üniversitenin işletme (n = 129), fizik (n = 82) ve psikolojik danışma ve rehberlik alanlarında (n = 98) öğrenim gören toplam 309 öğrenci (159 kız, 150 erkek) çalışmaya gönüllü olarak katılmıştır. Örneklemde yer alan öğrencilerin, sırasıyla, 33'ü 1. sınıflarda, 97'si ikinci sınıflarda, 84'ü üçüncü sınıflarda ve 95'i 4. sınıflarda öğrenim görmektedir. Örneklemde yer alan öğrencilerin yaşları 17 ile 33 arasında değişmektedir. Buna göre, öğrencilerin yaş ortalaması 21.21 standart sapması ise 1.83 olarak hesaplanmıştır. Öğrencilerin ilgi hissi olarak meraklarını ölçmek için Bilgi Kaynaklı Merak Ölçeği, yoksunluk hissi olarak meraklarının ve problem çözme ve enformasyonu arama davranışlarını motive eden gerilimin ölçülmesi içinse Yoksunluk Hissi Olarak Merak Ölçeği kullanılmıştır. Öğrencilerin başarı amaçlarına ilişkin veriler ise Başarı Amaçları Anketi aracılığıyla elde edilmiştir. Araştırmada kullanılan tüm ölçme araçları araştırmacı tarafından Türkçe'ye çevrilmiş ve Yabancı Diller alanından iki uzmanın yardımı aracılığıyla da tekrar-çeviri yöntemi uygulanmıştır. Ölçekler öğrencilere araştırmanın amacını ve maddelere ilişkin görüşlerin nasıl belirtileceğine ilişkin açıklamaları içeren birer yönergeyle verilmiştir. Ölçeklerin uygulanması esnasında öğrencilerden gelen sorular yanıtlanmıştır. Verilerin çözümlenmesinde tek örneklem t-testleri, korelasyon ve hiyerarşik regresyon analizleri kullanılmıştır.

Araştırmanın Bulguları: Araştırmada tek örneklem t-testi aracılığıyla elde edilen betimsel bulgular, öğrencilerin hem yoksunluk hem de ilgi hissi olarak merak düzeylerinin görece yüksek düzeylerde olduğunu göstermiştir. Öğrencilerin üniversitede öğrenim gören öğrenciler olduğu ve üniversite öğrenim sürecinin de merak düzeyini canlı tutan eğitimsel/öğretimsel uygulamalar içerdiği

dikkate alındığında, öğrencilerin bilgi kaynaklı merak düzeylerinin görece yüksek değerlerde olmasının beklenen bir bulgu olduğu söylenebilir. Bununla birlikte, tek örneklem t-testi sonuçları öğrencilerin performanstan kaçınma amaçlarından çok, performansa yaklaşma ve ustalaşma amaçlarını kabul etme eğilimlerinin yüksek olduğuna işaret etmektedir. Araştırmanın bu bulgusu, hakkında gittikçe artan miktarda kanıtın elde edildiği çoklu amaç çerçevesi ile uyumlu bir bulgudur. Nitekim tek örneklem t-testi aracılığıyla elde edilen bu bulgu, korelasyon analizi aracılığıyla elde edilen bulgularla da desteklenmiştir. Buna göre, korelasyon analizi sonuçları ilgi hissi olarak meraka göre yoksunluk hissi olarak merakın, öğrencilerin başarı amaçlarıyla anlamlı ve görece daha yüksek düzeyde ilişkili olduğunu gösteren anlamlı bir çerçeve sağlamıştır. Beklentinin aksine, hiyerarşik regresyon analizi sonuçları ilgi hissi olarak merakın değil, yoksunluk hissi olarak merakın öğrencilerin başarı amaçlarını yordamakta olduğunu göstermiştir. Daha açık bir ifadeyle, yoksunluk hissi olarak merak, ustalaşma ve performansa-yaklaşma amaçlarının yoksunluk hissi olarak merak güdümlü olduğunu gösterecek şekilde, öğrencilerin ustalaşma ve performansa-yaklaşma amaçlarını anlamlı düzeyde ve arka plan değişkenlerinin (yaş, cinsiyet, program türü) etkisinden bağımsız olarak yordamıştır. Araştırmadan elde edilen bulgular doğrultusunda eğitimsel çıkarsamalar ve gelecekte yapılabilecek araştırmalar üzerinde de durulmuştur.

Ancak, araştırmanın bulgularının görece küçük bir örneklem üzerinden ve tek bir üniversiteden elde edilmiş olması genellenebilirliklerini sınırlamaktadır. Bununla birlikte, araştırmanın kesitsel olması ve verilerin korelasyonel yöntemlerle incelenmesi, elde edilen bulguların nedensel bir biçimde yorumlanmasına olanak vermemektedir. Dolayısıyla, bu çalışmadan elde edilen bulgular gelecekte yapılacak olan araştırmalarda dikkatle yorumlanmalıdır.

Araştırmanın Sonuçları ve Önerileri: Sonuç olarak bu çalışma üniversite öğrencilerinin hem yoksunluk hissi olarak merak hem de ilgi hissi olarak merak düzeylerinin görece yüksek olduğunu göstermiştir. Bununla birlikte, çalışmadan elde edilen bulgular başarı amaçlarının çoklu amaç çerçevesi içerisinde ele alınmasının daha uygun bir yaklaşım olduğuna işaret edecek şekilde, öğrencilerin performanstan-kaçınma amaçlarından çok, hem performansa yaklaşma hem de ustalaşma amaçlarını benimseme eğiliminde olduklarını ortaya koymuştur. Son olarak, çalışmanın sonuçları bir yoksunluk hissi olarak merakın öğrencilerin hem ustalaşma hem de performansa-yaklaşma amaçlarının anlamlı bir yordayıcısı olduğunu göstermiştir. Söz konusu sonuçlar ışığında, öğrencilerin bilgi kaynaklı merak düzeylerinin gerek eğitim ortamlarında gerekse gelecekte yapılacak olan araştırmalarda dikkate alınmasının önemli doğurguları olacağı söylenebilir. Gelecekte farklı yaklaşımlarla ve daha büyük örneklemler aracılığıyla gerçekleştirilecek araştırmalar motivasyon ve bilgi kaynaklı merak konularına ilişkin olarak eğitim bilimsel anlamda önemli katkılar sağlayabilirler.

Anahtar Sözcükler: Bilgi kaynaklı merak, başarı amaçları, üniversite öğrencileri, hiyerarşik regresyon analizi

Does Being Leisure-Oriented Preclude a Person's Commitment to Work?

Şöheyda Göktürk*

Suggested Citation:

Göktürk, Ş. (2009). Does being leisure-oriented preclude a person's commitment to work? *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 145-160.

Abstract

Background/Problem Statement: Explaining the nature of the relationship between the domains of leisure and work has become important, as some studies show that today's workforce values leisure over work—or at least wants to balance leisure and work activities. Related literature has mostly focused on the influence of work on leisure. However, some research suggests that there are important leisure influences that operate on work. Very little attention has been paid to the direction of the influence between leisure and work. In addition, there are conflicting findings regarding whether gender has any role in this relationship.

Purpose of Study: The aim of this study is to examine whether teachers' attitudes toward leisure predict their work attitudes as indicated by their organizational commitment and commitment to their workgroup. Determining the role of gender in this relationship is also the aim of this study. The ultimate purpose is to increase our understanding of teachers' behavior at work and hence improve human resource management practices in schools.

Methods: A cross-sectional correlational survey design was used in this study. Participants were 610 primary and secondary school teachers selected with a cluster sampling technique from 65 schools in the province of Kocaeli, Turkey. The leisure ethic dimension of the Multidimensional Work Ethic Profile, as well as organizational commitment and workgroup attachment scales, were administered to collect data. Multiple regression analysis was applied to test the relationship between leisure attitudes (predictor variable) and work attitudes (criterion variables). Hierarchical regression analysis was applied to test the mediating effect of workgroup commitment on the leisure and organizational commitment relationship.

Findings/Results: The influence of leisure on workgroup and organizational commitments were positive and significant for men. For women, the same influences were negative but not statistically significant for workgroup commitment, and negative and statistically significant for organizational commitment. Further, hierarchical regression analysis showed that, for

* Asst. Prof. Dr., Kocaeli University, Faculty of Education, sdoyuran@postamatik.com

men, the influence of leisure on organizational commitment was mediated by workgroup commitment.

Conclusion and Recommendations: Valuing leisure is not harmful; in fact, it is beneficial for work—at least for the male teacher sample. High leisure orientation increases the workgroup commitment of male teachers, which in turn increases their commitment to the organization. School principals with creative human resource practices such as team sports can satisfy male teachers' leisure needs and at the same time increase their commitment to the organization through increased workgroup commitment. The negative influences of pro-leisure attitudes on female teachers' work commitment also deserve attention in human resource management.

Keywords: Leisure orientation, organizational commitment, teachers, work commitment, workgroup commitment, gender

The literature on leisure has mostly focused on the influence of work on non-work areas. However, some research suggests that there are important influences which operate from non-work activities to work activities (Cohen, 1997; Kirchmeyer, 1992; Kohn, 1990). The number of studies investigating the influence of non-work domains, such as family and community, on work domains is small; the amount of research related to the impact of leisure on work domains is even smaller (e.g., Kirchmeyer, 1992; Scholarios & Marks, 2004; Snir & Harpaz, 2002).

Throughout history, the importance of work for employees has been one of the dominant subjects of psychological investigations of occupational behavior (Pryor & Davies, 1989). However, an assumption that work rather than non-work is the only force affecting employees' organizational behavior can be misleading. Explaining the nature of the relationship between the domains of non-work and work has become important, since some studies show that time spent in leisure activities has been increasing, especially in developed countries (Harpaz, 1999). In addition, today's workforce values non-work over work or, at least, wants to balance non-work and work activities (Schein, 1996; Sturges & Guest, 2004). Therefore, organizations have been advised to adopt human resource management strategies that enable them to integrate leisure, family and the work of employees in order to increase favorable organizational outcomes such as organizational commitment (Quik, C., Henley & Quik, D., 2004; Scholarios & Marks, 2004; Sturges & Guest, 2004).

Some studies show that women's leisure is more related to their work than that of men (e.g., Kelly & Godbey, 1992). Kirchmeyer (1992), one of the few researchers who has studied non-work to work influences, found women and men to be almost identical in the extent to which they perceive non-work participation (i.e., community involvement and recreation) as enhancing and supporting of their work. Other findings, however, underscore the existence of negative influences from prevailing ideologies of motherhood and womanhood on women's sources of leisure (Kay, 2000). That is, women behaving in gender-stereotypical ways, such as having children and changing their work patterns to accommodate childcare considerations, experience more disadvantages than benefits from the practical and psychological attributes of work (Lane & Piercy, 2003).

Is a concern about leisure different than a specific concern to avoid any commitment to work? Are commitments to leisure and work necessarily competing opposites? Very

little attention has been paid to the direction of the influence between leisure and work, and there are conflicting findings regarding whether gender has any role in this relationship. These issues provide the motivation for the current study. More specifically, determining whether an employee's level of leisure orientation predicts his/her work commitment and the role of gender in this orientation are the main concerns of this study.

The remainder of this paper is constructed as follows. First, the theoretical framework and related literature are presented. Under this heading, the leisure to work relationship, variables of the study, and the hypotheses are expressed. The methods section gives the design of the study and describes the sample and procedure, data gathering instruments, and statistical analysis. Next, results of the study are offered. Finally, a discussion of the findings and limitations of the study are given, and recommendations for future research are stated.

Theoretical Framework and Related Literature

Leisure to work relationship. Three different models explaining the relationship between non-work and work are cited in the literature (e.g., Elizur, 1991; Judge & Watanabe, 1994; Near, Rice, & Hunt, 1980; Snir, 2002). Two of these models – “spillover” and “compensation” – were developed by Wilensky (1960, as cited in Elizur, 1991). The third, “segmentation,” was proposed by Dubin (1973, as cited in Elizur, 1991).

The spillover or generalization model states, “Attitudes and practices developed in one sphere of life can spill over into another – killing time at work can become killing time in leisure, apathy in the work place can become apathy in politics, alienation from one, alienation from the other” (Wilensky, 1960, p.545). According to the spillover model, an individual's non-work attitudes and behaviours should be positively related to his or her work-related attitudes or behaviours: An individual (dis)satisfied in one domain will be (dis)satisfied in the other (Elizur, 1991; Judge & Watanabe, 1994). The compensation hypothesis, on the other hand, proposes that the attitudes or behaviours in work and non-work domains are negatively correlated: Individuals experiencing deprecation at work seek compensatory reward in non-work (Elizur, 1991; Snir & Harpaz, 2002; Judge & Watanabe, 1994). The third model states, “...society's major institutions are basically independent of each other, and that they, therefore, impinge upon behaviour very differently” (Dubin, 1971, p.9). That is, work and non-work domains are independent of one another, and a non-significant correlation is found between work and non-work attitudes or behaviours (Dubin, 1958, 1973, as cited in Elizur, 1991).

The literature is not clear on which model is more influential in explaining the relationship between non-work and work in general and leisure and work in particular (e.g., Kirkcaldy, Trimpop, Fischer, & Furnham, 1997; Snir & Harpaz, 2002). Kirkcaldy and Cooper (1993) found that British managers were inclined to display negative spillover between their work and non-work domains. In another study with German managers, Kirkcaldy, Shephard and Cooper (1993) reported a compensatory link between work and leisure. These differences might be explained by Primeau's (1996) finding that the meaning of work and leisure is culturally determined. The present study, therefore, examines whether or not leisure attitudes or orientations of Turkish teachers predicts their work commitment, and therefore, tests the spillover model in a traditionally Muslim and, until recently, collectivist culture (e.g., Hofstede, 1980; Schwartz, 1994).

Work commitment. A review of the concept of work commitment in the literature revealed some thirty different definitions (Morrow, 1983). This construct is found to embody multiple factors (Blau, Paul, & St-John, 1993; Cohen, 1997; Morrow, 1983; Randall

& Cote, 1991); the universally accepted ones in the literature are work ethic, career and organizational commitments, workgroup attachment, and job involvement (Morrow, 1993; Randall & Cote, 1991). Work commitment factors are well-recognized work-related concepts that have been shown to predict important outcomes such as turnover, turnover intentions, performance, job satisfaction, absenteeism, and tardiness (Becker, 1992; Blau, 1986; Jaramillo, Mulki, Marshall, 2005). In this study, due to the limitation of the data, only workgroup and organizational commitment forms were investigated.

Lewin's field theory (1943), which is frequently applied to commitment research (e.g., Bentein, Stinglhamber, & Vandernberge, 2002; Mathieu, 1991; Mathieu & Hamel, 1989), is also helpful in explaining the direction of prediction among the variables being studied. Based on this theory, it is generally hypothesized that the most proximal or local entity in an employee's work environment should directly affect behavioural or attitudinal responses at work. Due to close physical and psychological proximity, local entities provide more control and feedback to one's work and therefore are more critical for orienting and guiding an employee's behaviour at work (Lawler, 1992; Mueller & Lawler, 1999). Another study by Lewin (1951) also suggested that "the influence of distal entities in the work environment on employee behaviour should be mediated by the proximal ones" (as cited in Bentein et al., 2002, p.345). Based on this rationale, in this study leisure attitude can be assumed to be the most proximal entity to an individual employee due to its intra-personal nature, whereas workgroups and the organization were the second and the distal entities, respectively. Following this line of thought, it is possible to hypothesize that a person's leisure attitude predicts his/her commitment to both workgroup and organization, and that workgroup commitment may mediate the influence of leisure orientation on organizational commitment.

Gender. Gender is cited as the most commonly used demographic variable in work attitude studies (Elizur, 2001). Studies related to the effect of gender on non-work (i.e., family, parenting and leisure) to work relationships do not show much agreement. Roehling, Jarvis, and Swope (2005) investigated negative family-work spillover among Black, White, and Hispanic Americans using a nationally representative sample. They found that Hispanic-Americans had the greatest disparity between men and women in negative family-to-work spillover, followed by White-Americans. Black-American men and women had similar levels of negative spillover. In another study, the effects of gender and academic rank on the incidence of spillover between work and non-work were assessed among university faculty in the United States (Sorcinelli & Near, 1986). Both work-family linkages and work-leisure linkages were examined. Negative spillover was found for this population, although no significant gender differences were found. Kirchmeyer (1992), on the other hand, found no difference between men and women in terms of community and leisure to work effects; both groups indicated that leisure and community participation enhanced, facilitated, and supported their work life.

Although the literature is limited and controversial, it may still be possible to conclude that differential ethnic and/or cultural backgrounds may interact with gender in the influence of leisure to work spillover. In this study, considering the unequal gender role perceptions that still pervade Turkish culture, even among professional women (Aycan, 2004), the positive influence of leisure orientation on work commitment is expected to be rather weak for women compared to men.

Statement of hypotheses. Based on the spillover model (Wilensky, 1960), Lewin's (1943) Field Theory, and the related literature, the following hypotheses were tested in this study. Hypothesis (1a): One's level of leisure orientation will predict his or her

workgroup commitment. Hypothesis (1b): One's level of leisure orientation will predict his or her organizational commitment. Hypothesis (2): The influence of leisure orientation on work commitment, as stated in hypotheses 1a and 1b, will be stronger for men than for women. Hypothesis (3): One's commitment to one's workgroup will mediate the relationship between leisure orientation and organizational commitment.

Method

Research Design

Because the study variables were not manipulatable by the researcher, the stated hypotheses search for predictive relationships using correlational and regression techniques, and both predictor and criterion variables were gathered at only one point in time, this study is in the form of a cross-sectional correlational survey.

Sample and Procedure

The sample in the present study consisted of elementary school teachers selected by using a cluster sampling technique from 65 schools in the province of Kocaeli, Turkey. That is, among a total of 342 public elementary schools in Kocaeli, 65 schools or clusters were randomly selected. First, selected schools' administrators were called and their consent for the research and determination of appropriate times to visit were requested. Then, selected schools were visited by the researcher and four of the students who became volunteers for data gathering in exchange for extra credit in the educational research methods course.

An explanation of the purpose of the study was given to available teachers at the time of visit in randomly selected schools (clusters), and they were invited to participate. Apart from a few (statistically negligible number of) teachers (attrition ranged from none to 2 by school), all agreed to take part. As a result, within each cluster a minimum of 4 and a maximum of 17 teachers were included the study. A total of 610 teachers were invited to participate in the study; 599 of them accepted and 594 of them were used. This represents quite a high response rate.

Participants were 51 % female. About 80 % had a four-year college degree. About 45 % were younger than 31, and 35% were between the ages of 32 and 45. About 70-75 % of the participants indicated that they had more than two years of organizational and occupational tenure. Finally, 68 % of the respondents were married, and 36% of those who were married had at least one pre-school age child.

In sum, in this study a combination of random and cluster selection techniques was used as a sampling strategy. The important point is that "The bias-free characteristic of random selection applies to the sample of larger units [in this study, schools], not to the elements [that is, teachers]" (Fox, 1969, p.342).

Instruments

Each participant completed a questionnaire composed of demographic background questions and instruments to measure two forms of work commitment and leisure orientation. Measures used in the study are standard scales adopted from the literature. The scale items were translated into Turkish first, than translated back into English by two bilingual speakers to ensure that the translation had the same meaning in both languages.

All constructs were measured on a 5-point scale (1 = strongly disagree or no opportunity to 5 = strongly agree or a lot of opportunity). In addition, the term

“organization” was changed to “school” in all measures. Other similar minor changes were applied as needed to adapt the scales to the participants’ occupation (i.e., teaching).

Leisure orientation. The 10-item leisure ethic dimension of the Multidimensional Work Ethic Profile (Miller, Woehr, & Hudspeth, 2002) was used to assess the leisure orientation of participants. For the current sample, the reliability and validity studies of this and the following measures were performed by the researcher. For these purposes, out of 594 valid cases, 105 male and 108 female teachers were randomly selected by the computer. Using these randomly selected 213 cases, internal consistency reliability and factor analysis procedures were conducted to assess the construct validity of the measure. The alpha coefficient for the scale was .83 for the total sample of 213. The scale was sufficiently reliable for men ($\alpha = .81$) and women ($\alpha = .79$) separately. In addition, factor analysis yielded one dimension with comparable loadings for both sexes, ranging from a minimum of .43 and .38 to a maximum of .83 and .79 for men and women, respectively. This demonstrates that the measure has good construct validity for the current sample. Miller et al. (2002) defined leisure orientation as pro-leisure attitudes and beliefs about the importance of non-work activities. Sample items included, “People should have more leisure time to spend in relaxation,” and “Life would be more meaningful if we had more leisure time.”

Organizational commitment. Mowday, Porter, and Steers’s (1979) 9-item short form measure was used to assess organizational commitment. Mowday et al.’s definition of the concept resembles the affective dimension of Meyer and Allen’s (1990) three dimensional commitment conceptualization. Using the 213 randomly selected cases mentioned above, reliability and validity studies of this scale were performed for the current study. The alpha coefficient for the scale was .83 for the total sample of 213. The scale was also sufficiently reliable for men ($\alpha = .85$) and women ($\alpha = .83$) separately. In addition, factor analysis yielded one dimension with comparable loadings for both sexes ranging from a minimum of .45 and .42 to a maximum of .87 and .84 for men and women, respectively. This demonstrates that the measure has good construct validity for the current sample. Sample items included, “I am proud to tell others that I am part of this organization,” and, “I talk up this organization to my friends as a great organization to work for.”

Workgroup Commitment. Sheldon’s (1971, as cited in Randall & Cote, 1991) three-item social involvement scale was expanded into a six-item workgroup attachment scale by Randall and Cote (1991). In this study Randall and Cote’s (1991) six-item workgroup attachment scale was used to operationalize the concept of workgroup commitment. The current study uses the concept as defined by Randall and Cote: “An employee’s sense of faithfulness to and collaboration with others in the organization.” Sample items included, “Some of my best friends are the people I work with,” and “I feel very much part of the people I work with.” The randomly selected 213 cases of male and female teachers discussed above were also used for the reliability and factor analysis of this scale by the researcher. The scale was satisfactorily reliable for both men ($\alpha = .86$) and women ($\alpha = .83$). Results of factor analysis yielded one factor for both male and female respondents; factor loadings ranged from a low of .44 to a high of .89. The latter demonstrates that the measure has good construct validity for the current sample.

Demographic Variables. In this study, age and occupational and organizational tenure were measured as ratio scales, gender (0 = women, 1 = men) and having a preschool age child at home (0 = no, 1 = yes) as dichotomous variables, and education on a scale from 1 to 5 (1 = two-year college, 5 = doctoral degree). These variables are frequently cited in the literature as correlates of work attitudes in general, and work commitments in particular

(e.g., Cohen, 1999). Salary was excluded from the analysis because all teachers in public schools in Turkey are paid in accordance with a single payment rate policy.

Analysis

Before proceeding with further analysis, normality and linearity of data were checked as a preliminary analysis by using SPSS version 10. In addition, a multicollinearity test was applied among independent variables. To detect any differences between men and women with respect to independent variables, *t* tests and chi-square tests were also applied. Finally, bivariate and partial correlation analyses were run among research variables to obtain a clearer picture of how one variable relates to another. In order to test the first hypothesis, hierarchical regression analysis was run for each dependent variable. For the second hypothesis, the same procedure was applied to male and female samples separately by splitting the data.

Finally, Baron and Kenny's (1986, cited in Kenny, 2008) four-stage hierarchical regression analysis model was applied to test the mediating effect of workgroup commitment on the leisure and organizational commitment relationship. According to Baron and Kenny (1986, cited in Howell, n.d.) to claim a mediating relationship there must be

a significant relationship between the independent variable and the mediator; a significant relationship between the mediator and the dependent variable; a significant relationship between the independent and dependent variable; and when the mediator and the independent variable are entered simultaneously to predict the dependent variable, the previously significant path between the independent and dependent variables must be greatly reduced, if not non-significant. (p. 2)

Findings and Results

In the preliminary analysis, no violations of assumptions were detected in the data. Among predictor variables, age and occupational tenure were collinear. To eliminate collinearity, occupational tenure was removed from the regression analysis. Chi-square and *t* test results showed that men and women did not differ significantly with respect to any of the demographic variables except age. That is, the men in the sample were significantly older than the women (means of 35 and 30 years, respectively), $t = 5.03$, $p < .001$. In addition, pro-leisure attitude levels of men and women did not differ significantly.

Correlations

For the whole sample, the intercorrelations (r 's) among some predictor variables (age, occupational and organizational tenure and education level) yielded from high to moderate values. Correlation coefficients between age and occupational and organizational tenures and education levels were .82 ($p < .01$), .53 ($p < .01$), and -.56 ($p < .01$) respectively. That is, older teachers were more likely to have less education and more experience in their occupation and organization. In addition, correlation coefficients between pro-leisure attitudes and the other predictor variables were not significant.

Predictor variables, with the exception of education level ($r = -.15$), were positively correlated to criterion variables; r values ranged between .10 and .15 ($ps < .05$). One variable of interest, pro-leisure attitudes, was also positively correlated with workgroup and organizational commitment ($r = .11$ ($p < .01$) and $r = .05$ ($p > .05$), respectively). In

addition, the bivariate correlation between criterion variables—workgroup and organizational commitments—gave $r = .49$ ($p < .01$).

Since gender had significant relationships with the forms of work commitment studied and the researcher wanted to examine gender influence, a separate correlation analysis was run for female and male groups using the same variables. For the female sample, the relationship between leisure and forms of commitment was negative in direction but not significant (r ranges between about $-.04$ to $-.10$, $ps > .05$). For the male sample group, on the other hand, the relationships between leisure and workgroup and organizational commitments were positive and r values were $.33$ and $.24$ ($ps < .01$), respectively. Table 1 presents an overview of the mean scores, standard deviations, and correlation coefficients between the variables for men and women separately.

Table 1
Mean Scores, Standard Deviations, and Correlation Coefficients of All Variables for Men (Below the Diagonal, N = 292) and Women (Above the Diagonal, N = 302)

	Mean(sd)		r							
	M	W	1	2	3	4	5	6	7	8
1. Have children (0= No, 1= Yes)	%45	%42	-	.51**	-.34**	.47**	.38**	.04	-.02	-.06
2. Age	35.10 (8.7)	30.86 (4.4)	.49**	-	-.59**	.81**	.55**	.11	.13*	.22**
3. Education	3.70 (.79)	3.74 (.77)	-.29**	-.53**	-	-.47**	.38**	-.07	-.13*	-.24**
4. Occupational Tenure	14.6 (7.9)	13.3 (7.1)	.45**	.83**	-.60**	-	.52**	.10	.15**	.20**
5. Organizational Tenure	6.12 (2.0)	5.43 (1.8)	.35**	.52**	-.33**	.53**	-	.20**	.11	.13*
6. Pro-leisure Attitude	2.94 (.69)	2.94 (.60)	.06	-.02	.01	-.04	.07	-	-.04	-.10
7. Workgroup Com.	3.31 (.65)	3.40 (.67)	.03	.05	-.10	.11	.24**	.33**	-	.43**
8. Organizational Com.	3.35 (.91)	3.43 (.88)	.08	.05	-.04	.06	.09	.24**	.54**	-

Note. * $p < .05$, ** $p < .01$ (two-tailed); M= Men, W= Women; Com.= Commitment.

Summarizing the correlation analysis results, the relationships between pro-leisure attitudes and workgroup and organizational commitment were positive in direction but only significant for the former variable for the whole sample. Separate correlation analyses for female and male groups on the same relationships showed that the forms of commitment made by females were negatively but not significantly correlated to a pro-leisure orientation. In contrast with the females, a pro-leisure attitude on the part of males was positively and significantly correlated with workgroup and organizational commitment.

For the male sample, the existence of positive and significant relationships between leisure and organizational commitment, leisure and workgroup commitment, and workgroup and organizational commitment ($r = .54, p < .01$) might indeed be a prelude to the identification of a mediating effect of workgroup commitment on the relationship between leisure and organizational commitment, as hypothesized previously.

Results of Hierarchical Regression Analysis

The results of Hypotheses 1a and 1b can broadly be deduced from the preliminary analysis above. However, to see the net effect of leisure attitude on workgroup and organizational commitment above and beyond the influence of demographic variables, hierarchical regression analysis with demographic variables in the first step and pro-leisure attitudes in the second step was first applied for the whole (total) sample to test hypotheses 1a and 1b. The contribution of pro-leisure attitude to the prediction of the dependent variable workgroup commitment was found to be small but significant ($B = .10, SE B = .42, p < .05$); on the other hand, pro-leisure attitude did not predict organizational commitment significantly. With these results hypothesis 1a was accepted and hypothesis 1b was rejected.

A similar analysis was applied separately to the male and female samples to test hypothesis 2. The results presented in Table 2 indicate that for the male sample, about 5 and 7 % of the total variance in organizational commitment and in workgroup commitment scores, respectively were explained by the level of pro-leisure attitude above and beyond the demographic variables controlled for in this study ($\Delta F = .000$). The ANOVA table indicated that the model as a whole was significant for workgroup commitment [$F(5, 287) = 12.85, p < .0001$]; for organizational commitment, only Model 2 was significant [$F(5, 287) = 4.42, p < .005$]. For the female sample, on the other hand, pro-leisure attitude explained 1.5 % of the variance in organizational commitment [$F(5, 297) = 6.45, p < .0001$]. Moreover, neither Model 1 (given in Step 1) nor Model 2 (given in Step 2) was significant for workgroup commitment for the female sample. Therefore, the results for workgroup commitment for the female sample are not given in Table 2.

Since the correlation analysis showed that the female sample's pro-leisure attitude did not correlate significantly with either workgroup or organizational commitment, the mediating effect of workgroup commitment was tested only in the male sample. To satisfy Baron and Kenny's (1986, cited in Howell, n.d.) basic prerequisites discussed in the procedure section, using multiple hierarchical regression analyses the mediator variable (workgroup commitment) was regressed on the predictor variable (leisure attitude) first. Next, the criterion variable (organizational commitment) was regressed on the mediator variable (workgroup commitment). Then the criterion variable (organizational commitment) was regressed on the predictor variable (leisure attitude). Finally, the criterion variable was simultaneously regressed on the predictor and mediator variables. The demographic variables were entered into each regression on the first step as a simultaneous block. In Table 2, the results of the first and third steps are presented. Tables 3 and 4 present the results of the application of the second and final steps, respectively.

Table 2

Summary of Hierarchical Regression Analyses for Variables Predicting Organizational Commitment and Workgroup Commitment for Male (N = 292) and Female Samples (in the second row, N = 302)

Variable	B	R ²	ΔR ²
<i>Organizational Commitment</i>			
Step1		.001	
		.054**	
Age	-.007		
	.169		
Organizational tenure	.033		
	-.014		
Pre-schooler at home	-.046		
	-.166		
Education level	.009		
	-.155		
Step2			.048*
			.014*
Pro-Leisure Attitude	.217*		
	-.119*		
<i>Workgroup Commitment</i>			
Step 1		.057*	
Age	-.096		
Organizational tenure	.145**		
Pre-schooler at home	-.052		
Education level	-.048		
Step 2			.067*
Pro-Leisure Attitude	.257**		

Note: For workgroup commitment neither Model 1 (Step 1) nor Model 2 (Step 2) was significant for the female sample.

*P < .05, **P < .001.

Table 3

Summary of Hierarchical Regression Analyses for Control Variables and Workgroup Commitment Predicting Organizational Commitment for Male Sample (N = 292)

Variable	B	R ²	ΔR ²
<i>Organizational Commitment</i>			
Step1		.009	
Age	-.037		
Organizational tenure	-.070		
Pre-schooler at home	-.043		
Education level	-.001		
Step2			.304*
Workgroup Commitment	.574*		

Note: *P < .0001.

The results presented in Table 3 show that for the male sample, workgroup commitment has an important and significant impact on organizational commitment after controlling for the demographic variables.

Table 4

Summary of Hierarchical Regression Analyses for Control Variables, Leisure Orientation, and Workgroup Commitment Predicting Organizational Commitment for Male Sample (N= 292)

Variable	B	R ²	ΔR ²
<i>Organizational Commitment</i>			
Step1		.009	
Age	-.008		
Organizational tenure	.087		
Pre-schooler at home	-.063		
Education level	.008		
Step2			.306*
Pro-Leisure Attitude	.057		
Workgroup commitment	.549*		

Note: *P < .0001.

The results in Table 4 clearly indicate that the previously significant path between pro-leisure attitude and organizational commitment ($\beta = .22$; $p < .05$) presented in Table 2 became non-significant when the mediator, workgroup commitment, and the independent variable, pro-leisure attitude, were simultaneously entered into the regression to predict the dependent variable, organizational commitment. Therefore, the third hypothesis was confirmed for the male sample group: Workgroup commitment mediated the relationship between pro-leisure attitude and organizational commitment.

Conclusions and Recommendations

The results showed that although men and women were equally inclined to leisure, they did not reflect this onto their work equally in terms of organizational and workgroup commitment. Having a positive attitude towards leisure had positive spillover effects on men's organizational commitment through its positive influence on workgroup commitment. Consistent with Lewin's Field Theory, workgroup commitment positively influenced men's organizational commitment. The indirect positive influence of a pro-leisure attitude on organizational commitment through workgroup commitment for men is an encouraging finding for human resource managers in general and for school administrators in particular. Little research has examined the impact of human resource socialization practices on the development of different forms of commitment. Therefore the results of this study may point the way for researchers interested in studying such relationships. It may be possible to increase the workgroup commitment of male personnel through creating deliberate opportunities for socialization among workgroup members such as picnicking, partying, and team sports activities. This may help satisfy employees' leisure needs and at the same time increase their commitment to the organization through increased workgroup commitment.

The negative influences of pro-leisure attitudes on women's work commitment also deserve attention in human resource management. The gender differences in these relationships may be attributed to the differential role perceptions that exist

between men and women in Turkish culture (Fikret-Pasa, Kabasakal, & Bodur, 2001; Kagıtcıbası, 1986). That is, in Turkey, women are still seen as primarily responsible for the care of family, children and husband, and secondarily for work. The prevalence of such differential perceptions of gender roles, even among Turkish women and men with higher job positions (Aycan, 2004), might have resulted in female teachers experiencing a conflict over commitment to family and workgroup or organization. That is, female teachers might have perceived the time spent with(in) the workgroup or organization as a barrier to the intrinsic rewards or leisure that they would otherwise experience with family and children. This interpretation is in line with findings indicating an overlap between family and leisure satisfaction (Kelly, 1983). Since women in Turkey are still perceived as more responsible for the care of family, Kelly's findings might apply more to Turkish women than to men.

Overall, it is possible to state that the findings of this study support the spillover model rather than the segmentation model, which asserts that non-work and work domains are unrelated and no relation exists between the domains (Dubin, 1971). Nevertheless, for future research, it seems important to test some of the above inferences, such as possible gender differences in family and leisure satisfaction, especially in cultures where perceptions of gender role inequity still exist. In addition, it may be important for future research to test whether workgroup commitment mediates the relationship between pro-leisure attitudes and organizational commitment in non-collectivist cultures where individuals are motivated in their leisure activities by their need for self-determination rather than their need for relatedness (Caldwell, 2005; Triandis, 1996). In sum, it seems necessary to include leisure aspects of non-work domains in future research to clarify the questions posed above.

It is especially important to remember that in this current study cross-sectional data were used and therefore causal inferences should not be made concerning the various hypothesized relationships. Inferences made here about the direction of the relationships between leisure and work commitment were based on theory and the previous research rather than the design of the study. Future research is required to test the findings of this study with a longitudinal research design. In addition, the relationships between predictor and criterion variables were not as strong as expected. One of the reasons for this could be the relatively small mean value and small standard deviation of the pro-leisure orientation variable for the whole sample. Future research needs to take this issue into consideration in the sample selection process to identify stronger or clearer relationships.

References

- Allen N. J. & Meyer J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology*, 63, 1-18.
- Aycan, Z. (2004). Key success factors for women in management in Turkey. *Applied Psychology: An International Review*, 53, 453-477.
- Becker, T.E. (1992). Foci and bases of commitment: Are they distinctions worth making? *Academy of Management Journal*, 35, 232-244.

- Bentein, K., Stinglhamber, F., & Vandernberge, C. (2002). Organization-, supervisor-, and workgroup-directed commitments and citizenship behaviours: a comparison of models. *European Journal of Work and Organizational Psychology*, 11(3), 341-362.
- Blau, G. J. (1986). Job involvement and organizational commitment as interactive predictors of tardiness and absenteeism. *Journal of Management*, 12, 577-584.
- Blau, G., Paul, A. & St. John, N. (1993). On developing a general index of work commitment. *Journal of Vocational Behavior*, 42, 298-314.
- Caldwell, L., (2005). Self-construal in Context: a response to Walker, Deng, and Dieser. *Journal of Leisure Research*, 37, 106-116.
- Cohen, A. (1997). Non-work influences on withdrawal cognitions: an empirical examination of an overlooked issue. *Human Relations*, 50(12), 1511-1536.
- Cohen, A. (1999). Relationships among five forms of commitment: an empirical assessment. *Journal of Organizational Behavior*, 20, 285-308.
- Dubin, R. (1971). *Work and non-work: Institutional perspectives*. (ERIC Document Reproduction Service No. ED055201). Retrieved from ERIC database.
- Elizur, D. (1991). Work and nonwork relations: The conical structure of work and home life relationship. *Journal of Organizational Behavior*, 12(4), 313-322.
- Elizur, D., & Koslowsky, M. (2001). Work values and commitment. *International Journal of Manpower*, 22, 593-599.
- Fikret-Pasa, S., Kabasakal, H., & Bodur, M. (2001). Society, organizations, and leadership in Turkey. *Applied Psychology: An International Review*, 50(4), 559-590.
- Fox, D. J. (1969). *The research process in education*. New York: Holt, Rinehart and Winston, Inc.
- Furnham, A. (1982). The protestant work ethic and attitudes toward unemployment. *Journal of Occupational Psychology*, 55, 277-286.
- Harpaz, I. (1999). The transformation of work values in Israel. *Monthly Labor Review*, 122, 46-50.
- Hofstede, G. H. (1984). *Culture's consequences. International differences in work-related values*. Newbury Park, CA: Sage Publications.
- Howell, D.(n.d.). *Mediating and moderating relationships*. Retrieved from <http://www.uvm.edu/~dhowell/StatPages/Methods/MedModCorrection.pdf>
- Jaramillo, F., Mulki, J. P., & Marshall, G. W. (2005). A meta-analysis of the relationship between organizational commitment and salesperson job performance: 25 years of research. *Journal of Business Research*, 58(6), 705-714.
- Judge, T. A., & Watanabe, S. (1994). Individual differences in the nature of the relationship between job and life satisfaction. *Journal of Occupational and Organizational Psychology*, 67, 101-107.
- Kagiticbasi, C. (1986). Status of women in Turkey: Cross-cultural perspectives. *International Journal of Middle East Studies*, 18(4), 485-499.
- Kay, T. (2000). Leisure, gender and family: The influence of social policy. *Leisure Studies*, 19, 247-65.
- Kelly, J. (1983). *Leisure identities and interactions*. London: Allen and Unwin
- Kelly, J., & Godbey, G. (1992). *The sociology of leisure*. State College, PA: Venture Capital.
- Kenny, D. (2008). Mediation. Retrieved November 14, 2008, from <http://davidakenny.net/cm/mediate.htm>.
- Khon, M. (1990). Unresolved issues in work/personality. In K. Erikson, and S. Wallas (Eds.). *The nature of work*. New Haven: Yale University Press.

- Kirchmeyer, C. (1992). Perceptions of non-work to work spillover: Challenging the common view of conflict ridden domain relationships. *Basic and Applied Social Psychology*, 13(2), 231-249.
- Kirkcaldy, B.D., & Cooper, C.L. (1993). The relationship between work stress and leisure style: British and German managers. *Human Relations*, 46, 669-80.
- Kirkcaldy, B.D., Shephard, R.J., & Cooper, C.L. (1993). Relationships between type A behaviour, work and leisure. *Personality and Individual Differences*, 15, 69-74.
- Kirkcaldy, B. D., Trimpop, R. M., Fischer, C., & Furnham, A. (1997). Leisure and work beliefs of British senior managers. *Journal of Management Development*, 16(6), 392-404.
- Lane, N., & Piercy, N. F. (2003). The ethics of discrimination: Organizational mindsets and female employment disadvantage. *Journal of Business Ethics*, 44, 313-325.
- Lawler, E.J. (1992). Affective attachments to nested groups: A choice-process theory. *American Sociological Review*, 57, 327-339.
- Lewin, K. (1943). Defining the "field at a given time". *Psychological Review*, 50, 292-310.
- Mathieu, J.E. (1991). A cross-level nonrecursive model of the antecedents of organizational commitment and satisfaction. *Journal of Applied Psychology*, 76, 607-618.
- Mathieu, J.E., & Hamel, K. (1989). A causal model of the antecedents of organizational commitment among professionals and nonprofessionals. *Journal of Vocational Behavior*, 34, 299-317.
- Miller, M. J., Woehr, D. J., & Hudspeth, N. (2002). The meaning and measurement of work ethic: Construction and initial validation of a multidimensional inventory. *Journal of Vocational Behavior*, 60, 451-489.
- Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior*, 14, 224-247.
- Mueller, C.W., & Lawler, E.J. (1999). Commitment to nested organizational units: Some basic principles and preliminary findings. *Social Psychology Quarterly*, 62, 325-346.
- Morrow, P. C. (1983). Concept redundancy in organizational research: The case of work commitment. *Academy of Management Review*, 8, 486-500.
- Morrow, P. C. (1993). *The theory and measurement of work commitment*. Greenwich, CT: JAI Press Inc.
- Mueller, C.W., & Lawler, E.J. (1999). Commitment to nested organizational units: Some basic principles and preliminary findings. *Social Psychology Quarterly*, 62, 325-346.
- Near, J., P., Rice, R. W., & Hunt, R. G. (1980). The relationship between work and nonwork domains: A review of empirical research. *The Academy of Management Review*, 5(3), 415-429.
- Petty, G. C. & Hill, R. B. (1994). Are women and men different? A study of the occupational work ethic. *Journal of Vocational Education Research*, 19(1), 71-89.
- Primeau, L., A. (1996). Work and leisure: Trencending the dichotomy. *American Journal of Occupational Therapy*, 50(7), 569-577.
- Pryor, R. G. L., & Davies, R. (1989). A comparison of conceptualizations of work centrality. *Journal of Organizational Behavior*, 10(3), 283-289.
- Quick, J. D., Henley, A., & Quick, J. C. (2004). The balancing act: At work and at home. *Organizational Dynamics*, 33(4), 426-438.
- Randall, D. M. & Cote, J. A. (1991). Interrelationships of work commitment constructs. *Work and Occupation*, 18, 194-211.
- Roehling, V. P., Jarvis, H. L., & Swope, E. H. (2005). Variations in negative work-family spillover among white, black, and hispanic American men and women. Does ethnicity matter? *Journal of Family Issues*, 26(6), 840-865.

- Schein, E. (1996). Career anchors revisited: implications for career development in the 21st century. *Academy of Management Executive*, 10(4), 80-88.
- Scholarios, D., & Marks, A. (2004). Work-life balance and the software worker. *Human Resource Management Journal*, 14(2), 54-75.
- Schwartz, S. H. (1994). Beyond individualism/collectivism: New cultural dimensions of values. In U. Kim, H. C. Triandis, Ç. Kağıtçıbaşı, S. C. Choi, and G.Yoon (Eds.). *Individualism and collectivism: Theory, method, and applications*. Thousand Oaks: Sage Publications.
- Snir, N., & Harpaz, I. (2002). Work-leisure relations: Leisure orientation and the meaning of work. *Journal of Leisure Research*, 34(2), 178-203.
- Sorcinelli, M. D., & Near, J. (1989). Relations between work and life away from work among university faculty. *Journal of Higher Education*, 60(1), 59-81.
- Sturges, J., & Guest, D. (2004). Working to live or living to work? Work/life balance early in the career. *Human Resource Management Journal*, 14(4), 5-20.
- Triandis, H. G. (1996). The psychological measurement of cultural syndromes. *American Psychologist*, 51, 407-415.
- Wilensky, H. L. (1960). Work, careers and social integration. *International Social Science Journal*, 12(4), 543-560. Retrieved from <http://unesdoc.unesco.org/images/0001/000168/016832eo.pdf>

Serbest Zaman Yönelimli Olmak Çalışanın İşe Bağlılığını Engeller mi?

(Özet)

Problem Durumu: Serbest zaman ve iş alanları arasındaki ilişkinin doğasını açıklamak özellikle günümüz işgücünün serbest zamanı işe ayrılan zamana göre daha değerli gördüğüne ya da en azından bu iki alandaki etkinliklerin insan yaşamında dengeli dağılmasının gerekliliğine inandıklarına dair araştırma sonuçlarının ortaya çıkmasıyla önemli hale gelmiştir. Bu alandaki literatür daha çok işin serbest zamana olan etkisine odaklanmıştır. Halbuki ilgili alandaki bazı önemli bulgular serbest zaman algısının iş algısına etkisinin olduğu yönündedir. Serbest zaman ve iş arasındaki ilişkinin yönüne ilişkin akademik ilgi çok az olup yine bu ilişkide cinsiyetin rolüne dair bulgular çelişkilidir.

Araştırmanın Amacı: Bu araştırmanın amacı öğretmenlerin serbest zaman tutumlarının örgütsel bağlılık ve çalışma grubuna bağlılık değişkenleriyle tanımlanmış olan iş tutumlarını anlamlı olarak yordayıp yordamadığını saptamaktır. Ayrıca, cinsiyetin serbest zaman ve iş tutumu değişkenleri arasındaki ilişkide modaratör değişken olup olmadığı da bu çalışmada incelenmiştir. Araştırmanın genel amacı ise öğretmenlerin iş yerindeki davranışlarını açıklamaya yönelik alanyazına ve okullardaki insan kaynakları yönetimi uygulamalarına katkı sağlamaktır.

Araştırmanın Yöntemi: Araştırma kesitli ilişkisel tarama modelindedir. Katılımcılar küme örnekleme yöntemiyle Kocaeli ilindeki 65 okuldan seçilmiştir. Toplam 610 ilköğretim okulu öğretmeni araştırmanın örneklemini oluşturmaktadır. Geçerlik ve güvenilirliği örneklem grubu için sağlanmış Çokboyutlu İş Etiği Profili ölçeğinin Serbest Zaman Etiği altölçeği, Örgütsel Bağlılık Ölçeği ve Çalışma-

grubuna Bağlılık Ölçeği sırasıyla serbest zaman tutumu, örgütsel ve çalışma-grubuna bağlılıkları ölçmek için uygulanmıştır. Serbest zaman tutumu ve iş tutumu arasındaki ilişkiyi analiz etmek için çoklu regrasyon analizi, çalışma-grubuna bağlılığın serbest zaman yönelim düzeyi ve örgütsel bağlılık arasındaki ilişkide arabulucu rolünü test etmek için de dört basamaklı hiyerarşik regrasyon modeli kullanılmıştır.

Bulgular: Serbest zamanın anlamı ve serbest zaman yönelim düzeyleri sırasıyla faktör analizi ve t-testi analizleri ile incelenmiştir; sonuçlar cinsiyete göre farklılaşma göstermemiş olmasına rağmen, erkeklerde, serbest zaman yönelim düzeyinin çalışma grubuna ve örgüte bağlılık düzeyini anlamlı olarak yordadığı ve yordayıcı ile kriter değişkenler arasındaki ilişkinin pozitif yönde olduğu bulunmuştur. Bayanlarda aynı yordayıcının kriter değişkenlerle ilişkisinin negatif yönde ve sadece örgütsel bağlılık değişkeniyle anlamlı olduğu bulunmuştur. Ayrıca, hiyerarşik regrasyon analiz sonuçları, erkeklerde çalışma grubuna bağlılık değişkeninin serbest zaman yönelim düzeyinin örgütsel bağlılığı kestiriminde yönlendirici (mediator) olduğunu göstermiştir.

Sonuç ve Öneriler: Serbest zaman ve iş birbirinden ayrı iki alan olarak görülmemelidir. Serbest zamana değer vermek işe verilen değeri azaltmamakta tam tersine işe olan bağlılığı artırmaktadır. Serbest zamana değer veren erkek öğretmenlerin çalışma grubuna bağlılığının da yüksek olma eğilimi gösterdiği ve çalışma grubuna olan bağlılığın dolaylı olarak örgüte olan duygusal bağlılığı artırdığı görülmüştür. En azından erkek öğretmen grubu için bu araştırmanın sonuçları bunu göstermiştir. Okul yöneticileri takım sporları ve piknik gibi yaratıcı insan kaynakları uygulamaları ile serbest zamana değer verme eğiliminde olan erkek öğretmenlerin serbest zaman ihtiyaçlarını çalışma gruplarında geçirilen kaliteli birlikteliklerle doyurabilir ve böylelikle onların çalışma grubuna bağlılık düzeylerini artırabilirler. Bu da çalışanların örgüte bağlılıklarını dolaylı olarak artırabilir. Kısacası serbest zamana değer veren erkek öğretmenler işyerindeki arkadaş gruplarına olan güçlü bağlılıkları- onlarla geçirdikleri hoşça zaman ve dostlukları aracılığı ile çalıştıkları örgüte de bağlanacaklardır. Serbest zaman yöneliminin bayan öğretmenlerin işe bağlılık düzeyleriyle genel olarak olumsuz ilişkisi ise insan kaynakları yönetimi alanındaki araştırmacılar için ilgiye değer bir bulgu olarak görülmektedir. Türkiye’de hala geçerliliğini koruyan geleneksel kadın rolünün gereği olarak serbest zamana değer veren bayan öğretmenlerin bu ihtiyaçlarını çalışma grubundaki birliktelik ve arkadaşlıklar yerine evde çocuk ve ailesiyle birlikte olarak doyurma eğiliminde olabilecekleri tahmin edilebilir. Bu konun gelecek araştırmalarda incelenmesi önerilebilir. Son olarak, bu araştırma, deseni açısından boylamsal bir araştırma olmadığı ve incelenen değişkenler arasındaki ilişkinin yönünü belirlemede teori ve önceki araştırmalara dayanıldığı için gelecek araştırmacıların bu araştırmayı boylamsal araştırma deseni ile tekrar etmesi önerilebilir.

Anahtar Sözcükler: Serbest zaman yönelimi, örgütsel bağlılık, işe bağlılık, çalışma grubuna bağlılık, öğretmenler

The Barriers Hindering Female Teachers From Becoming School Principals (Sample from Mersin Province-Turkey)

Yusuf İnandı*

Suggested Citation:

İnandı,Y.(2009). The barriers hindering female teachers from becoming school principals. *Eğitim Araştırmaları - Eurasian Journal of Educational Research*, 36, 161-178

Abstract

Problem Statement: Despite many positive developments, in practice it is still not possible for women to be equal to men in the twenty-first century. One of the fields in which inequality between men and women continues to exist is education. In educational institutions, women are not given the same opportunities as men to be promoted to higher ranks and senior positions which include school principals. In many research studies, the barriers preventing women from career advancement have been found to be the existence of a glass ceiling; stereotypes about the sexes; obstacles resulting from societal structure; barriers resulting from family and women's point of view towards administration; barriers resulting from education, working hours, age and marital status; and career obstacles resulting from school organization and environment. This paper examines the barriers preventing female teachers working in elementary schools from becoming school principals.

Purpose of Study: The aim of the research is to determine the barriers preventing female teachers working in elementary schools from becoming school principals by analyzing the views of teachers and school principals who work in public elementary schools.

Methods: A general scanning model is used in this descriptive research study. It was decided to use the simple random sampling approach, and 30 school principals and 202 teachers in the Mersin Province were randomly selected to constitute the sample for the research study. The researcher collected the data for the research by implementing a rating scale. By means of SPSS 11, the views of school principals and teachers about the career barriers preventing female teachers working in elementary schools from becoming school principals were determined.

Findings and Results: According to the findings, the position variable and the sex variable cause meaningful differences between the points of view of the participants. There are meaningful differences between the points of

* Assist.Prof.Dr., Mersin University, Faculty of Education, Dept. of Educational Science, Yenışehir / Mersin / Turkey, inandiyusuf@gmail.com

view of the participants about the dimensions of working hours, age, and marital status in terms of the variable of education level. Furthermore, there were meaningful differences between the points of view of the participants about societal sex stereotypes.

Conclusions and Recommendations: It can be concluded that the glass ceiling produced by society to discriminate against women continues to exist. The men in today's societies prioritize women's duties as motherhood, and being a housewife and a spouse. In contrast, the activities of career development are perceived as being men's work. It should be considered that female teachers, along with other individuals who constitute an organization, have an important responsibility to help female teachers to overcome the barriers to career development.

Keywords: Career advancement, gender stereotypes, school principals, teacher, barriers

Throughout the centuries, human beings have fought for ideals such as freedom and democracy. Due to their struggles and their devotion, these ideals have been realized. However, some ideals have not yet been fully realized and the struggle for them is still ongoing. One of these struggles is related to the issue of equality between men and women. After the industrial revolution, women started to gain new rights; a process which continued at an increasing rate in the twentieth century. For example, women obtained the right to vote and to possess property in many countries in the twentieth century; and the constitutions of many countries incorporated equality between men and women. Despite all of these positive developments, in practice women are still not equal to men in the twenty-first century. One of the fields in which the inequality between men and women continues to exist is education. In educational institutions, women are not given the same opportunities as men to be promoted to higher ranks and senior positions which include school principals. Therefore, the number of women in administrative positions is not in proportion to the number of women joining the labour force (Moreau, Osgood & Halsall, 2007; Çelikten, 2004; Athanassoula-Reppa & Koutouzis, 2002; Evetts, 2000).

There is also discrimination against women in the educational field in Turkey (Kadının Statüsü ve Sorunları Genel Müdürlüğü [KSSGM], 2000). According to Çelikten (2004) the number of women holding administrative positions in educational management has not increased in Turkey. The barriers preventing women from progressing in their career at the elementary school level show a parallelism between Turkey and the rest of the world. These barriers have been identified in numerous research studies (Evetts, 2000; Çelikten, 2004; Thompson, 2003; Zafarullah, 2000) and they are explained below:

Glass ceiling. The concept of a 'glass ceiling' which results from people's attitudes and the prejudice of organizations is defined as an invisible and insurmountable barrier that exists between women and senior management positions; and that prevents women from making progress (Çelikten, 2005; Özbilgin & Woodward, 2004; ILO, 1997). Throughout the world, although there is an increase in the rate of women joining the workforce and taking up a career in education, most women face

occupational discrimination in business life; and this ceiling prevents them from occupying senior administrative and professional positions (Moreau et al., 2007; Thompson, 2007).

Stereotypes about sexes, barriers resulting from societal sexes and societal structure. The prejudice of societies and individuals, in addition to social norms about women taking on senior level positions, create barriers for women. Stereotypes about the sexes are a foundation for the prejudice preventing women from being promoted to administrative positions. Education is one of the fields in which these stereotypes are seen to be prevalent (Usluer, 2000). Otaran, Sayın, Güven, Gürkaynak, and Atakul (2003) found that, although 56 % of teachers are female in elementary schools in Turkey, female teachers make up less than 3 % of administrators. This finding shows that, while the teaching profession is regarded as a very suitable job for women in Turkey, educational administration is perceived as a man's job.

The problem of women's employment in the administrative field in educational organizations is not only peculiar to Turkey. Similar situations of discrimination against women exist in many European countries including the Netherlands, Greece, and Italy. In these countries, the number of female teachers is greater than the the number of male teachers. Nonetheless, the number of female school principals is much less than the number of male school principals. For example, 13% of the school principals in the Netherlands, 41% of the school principals in Greece, and 46% of the school principals in Italy are female. When we compare Turkey with European countries, it is obvious that Turkey, in which the proportion of female school principals is less than 3%, remains behind these European countries (Athanasoula-Reppa, A. & Koutouzis, 2002, ¶ 41).

Barriers resulting from family. Societal beliefs, such as "the women's first priority is motherhood and becoming a good spouse", which exist in both the cultural and the structural dimension of career development and which are defined as role traps for women, negatively influence women's chances of becoming administrators. In other words, as Oplatka (2001) states, women are culturally expected to be caring, subjective, and personal. In addition, these societal beliefs are reproduced as barriers to career development and they also help to preserve these barriers (Evetts, 2000; Rimmer & Rimmer, 1997; Thompson, 2003). According to a research study carried out by the TUSİAD in Turkey (quoted by Çelikten, 2004), women who accept the traditional roles of motherhood and becoming a good spouse regard their role in the public field as having secondary importance; and they think that becoming a teacher is more appropriate than becoming a school principal.

Women's point of view about administration. Another important barrier is women's lack of self-confidence. It is an obvious fact that the women in administrative position are always pushed into this position by others. The attention that women continuously pay to other people's approval shows that they don't have confidence in themselves. This situation exists in the action dimension that Evetts (2000) defines. Women's lack of self-confidence and the Turkish culture cause them to be perceived as passive characters in Turkish society. The influence of the male hegemony on

women throughout their life results in women becoming passive individuals in society (Merle, 1999; Zafarullah, 2000).

Barriers resulting from education, working hours, age, and marital status. Education, working hours, age, and marital status play an important role in the process of women's career development. The fact that a woman in an administrative role has to work longer working hours than teachers causes most of them to reject becoming administrators; and the priority of bringing up their children causes them to become interested in administration later in life. Moreover, their efforts to advance their education and obtain certificates are indications that they think that they are inadequate in this field, that they lack confidence, and therefore try to create self-confidence by means of these documents (Çelikten, 2004; Hudson & Rea, 1998).

Career barriers resulting from school and the environment. The societal norms which emphasize that men have to be school principals are reproduced and protected in formal organizations such as schools (Thomson, 2003). In the institutions where administrative and organizational culture are firmly established in the patriarchal structure (Evetts, 2000), women's decisions about career development are negatively influenced. Moreover, the societal norms and beliefs emphasizing that men should be administrators have a great influence on families. Families prefer male school principals, or female school principals behaving like men. This is because the parents want to control their children and they consider that a serious atmosphere is created by authoritarian school principals, or by the female school principals who adopt an authoritarian approach like their male counterparts (Smulyan, 2000).

The main aim of this research is to determine the barriers preventing female teachers working in elementary schools from becoming school principals by examining the views of teachers and school principals who work in formal elementary schools. Subsequent to this main aim, various sub-aims were identified and the additional questions given below were addressed:

According to the views of teachers and school principals who work in formal elementary schools;

a-In terms of the position variable, is there a meaningful difference among the barriers preventing female teachers working in elementary schools from becoming school principals?

b-In terms of the sex variable, is there a meaningful difference among the barriers preventing female teachers working in elementary schools from becoming school principals?

c-In terms of the education variable, is there a meaningful difference among the barriers preventing female teachers working in elementary schools from becoming school principals?

d-In terms of the spouse's job variable, is there a meaningful difference among the barriers preventing female teachers working in elementary schools from becoming school principals?

Method

Research Design

A general scanning model is used in this descriptive research study. By means of this method, women's career barriers are identified according to the points of view of administrators and teachers. In addition, an attempt is made to describe the existing situation.

Population and Sample

The population of the research includes 150 school principals and 4200 teachers working in the elementary schools of Mersin Province in the period 2006-2007 (MEB, 2006). 41 out of 150 school principals were contacted; however, as some of the school principals didn't fill in the questionnaires and some of the questionnaires were not appropriately completed, 11 questionnaires are not included in the scope of research. 225 out of 4200 teachers were contacted, however, as 23 of the teachers did not fill in the questionnaire accurately, they are not included in the scope of research. As a sampling approach, simple random sampling was used in this research study because this sampling technique ensured that all of the participants in the population had an equal chance to be selected for the study. 30 school principals and 202 teachers were randomly selected from schools in the Mersin Province, and 232 school principals and teachers constituted the sample for the research. The sample for the school principals represents 20% of the total population of school principals. The sample for the teachers represents approximately 6% of the total population of teachers. 65% of the teachers were women and 35% of the teachers were men. 86% of the school principals were men and 14% of the school principals were women. 16% of the teachers and school principals had associate degrees; 75% of them had bachelor's degrees; 5% of them had master's degrees; and 4% of them were other graduates. In terms of marital status: 52% of the participants were married to teachers; 4% of the participants were married to retired people; 8% of the participants were married to civil servants; 12% of the participants were married to housewives; 6% of the participants were married to nurses; 4% of the participants were married to engineers; and 14% were married to people with other jobs.

Research Instruments

Rating scale. The researcher collected the data for the research by implementing a rating scale. The researcher developed the rating scale by reading the local and foreign literature. The first part of the rating scale includes questions about personal information. The second part of the questionnaire includes 42 questions related to female teachers' career obstacles.

The validity and reliability of the research instruments

For the validity of the research, the researchers consulted the opinion of experts in the field in order to have a scope of validity for the rating scale. The rating scale including 50 items were prepared for a pilot study. After the pilot study, explanatory factor analysis was made to see the structure validity of the rating scale; and for each factor, both sub and upper factor load analyses are given. The rating scale was

developed into a structure having six components and two factors. These two factors explained 56% of the total variance. Factor loads according to the sub-aims were analyzed: the sub factor loads for career barriers related to family are .44 and .82; the sub factor loads for the career barriers resulting from school and environment are .41 and .77; the sub factor loads for the barriers resulting from education, working hours, age, and marital status are .41 and .70; the sub factor loads for the obstacles resulting from societal sexes and societal structure are .52 and .74; the sub factor loads for the women's point of view about the administration are .48 and .69; and the sub factor loads for the barriers resulting from the economical problems are .73 and .77. Moreover, the total score was calculated and its relationship to each item has been analyzed in order to find out the item validity. The items for which the total correlation was much more than .40 are used in the scale; and 8 items for which the total correlation was less than .40 were not used. The scale used comprised 42 items.

For every dimension of the scale and for the bases of all items, reliability was checked and the Cronbach Alpha was found to be .94 for the entire scale. The Cronbach Alpha for the career barriers related to family sub-aim is .87; the Cronbach Alpha for the barriers resulting from school and environment sub-aim is .86; the Cronbach Alpha for the barriers resulting from education, working hours, age, and marital status is .87; the Cronbach Alpha for the barriers resulting from societal sexes and societal structure sub-aim is .83; the Cronbach Alpha for the women's point of view about the administration sub-aim is .81; and the Cronbach Alpha for the barriers resulting from the economical problems sub-aim is .84.

To check whether the sample was appropriate, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy Test was applied and this value was found to be .88. This value shows that the sample is very appropriate. To see whether the distribution was normal, the values of Barlett's Test of Sphericity were found to be χ^2 : 5209,509; df: 861; $p < .000$, and the value of Chi-Square was found to be 5209,509; df 861 ve $p < .000$. According to these results, multi-variable normal distribution was obtained. After using the scale, the analysis of the data was quantified. As a five-degree scale was used, the interval coefficients were calculated for four intervals ($5-1=4$) in the five-degree scale like this:

$$(4/5=0.80)$$

I never agree = 1.00-1.79,

I don't agree = 1.80-2.59

I am indecisive= 2.60-3.39

I agree= 3.40-4.19

I agree completely = 4.20-5.00

Procedure

The rating scale was applied to 225 teachers and 41 school principals. 223 teachers and 30 school principals filled in the scale appropriately and submitted it to the researcher.

Data Analyses

The data collected with the scale was inputted into the computer according to codes that were given to each question on the scale. When process of inputting the data was finished, the data was processed and analyzed. In the analysis of the data, SPSS 11 computer programme was used. By means of this programme, the views of school principals and teachers about female teachers' career barriers were determined using the values of the arithmetic mean and standard deviation. The T-test was applied for the independent groups to determine whether there was a meaningful difference among the views of the subjects according to the position and sex variables. One-way variance analysis was implemented to determine whether there was a meaningful difference among the views of teachers and school principals according to the variables of the level of education, marital status, spouse's job, and the number of children. For the cases in which differences were detected, the Tukey HSD test was applied. Conclusions and recommendations were made based on the findings and results. The level of meaningfulness was found to be $p < 0.05$ with $p < 0.01$.

Findings and Results

In the sub-aims part of this paper, the first sub-aim of the research was expressed as: "According to the position variable, is there a meaningful difference among the barriers preventing the female teachers working in elementary schools from becoming school principals?" In Table 1, the T-test and the results related to the first sub-aim of the research are presented. According to this result, the position variable causes a meaningful difference in societal sex stereotypes. According to the position variable, there are not any meaningful differences in the other dimensions between the views of school principals and teachers.

Table 1

T-test Related with the Views of Participants in Terms of the Position Variable

	Position	N	Mean	S	t	P
Resulting from family	Teacher	202	3,2605	,87325	,309	,758
	Administrator	30	3,2083	,79758		
School and environment	Teacher	202	2,4092	,77844	,985	,326
	Administrator	30	2,2630	,60962		
Working hour, Age and Marital Status	Teacher	202	2,8252	,77614	,708	,480
	Administrator	30	2,7200	,63918		
Societal Sex	Teacher	202	2,8738	,86540	2,357	,019*
	Administrator	30	2,4778	,81029		
The women's opinion about administration	Teacher	202	2,6351	,76487	-,020	,984
	Administrator	30	2,6381	,73491		
Economic Reasons	Teacher	202	2,7203	1,04011	,432	,666
	Administrator	30	2,6333	,94626		

* $p < 0.05$

The school principals' points of view about the barriers preventing the female teachers working in elementary schools from becoming school principals due to women's societal sex stereotypes is (Mean=2,4778); and the teachers' point of view about the the barriers preventing the female teachers working in elementary schools from becoming school principals due to women's societal sex stereotypes is (Mean=2,8738). When the school principals' points of view about this issue are compared with the teachers' points of view, it can be observed that there is a meaningful difference between the points of view of teachers and school principals ($t:-2,357;p<.019$). The teachers' perceptions are what cause the difference. Compared with school principals, more teachers think that societal sex stereotypes cause the barriers preventing the female teachers working in elementary schools from becoming school principals.

The second sub-aim of the research was expressed as: "According to the sex variable, is there a meaningful difference among the barriers preventing the female teachers working in elementary schools from becoming school principals?" In Table 2, in terms of the sex variable, the T test and its results related to the second sub-aim of the research are presented. According to these results, the sex variable causes a meaningful difference in the barriers resulting from family and societal sex stereotypes. In terms of the sex variable, a meaningful difference was not found among the views of teachers and school principals related to the influence of school and environment, economic reasons, women's opinion about the administration, working hours, age, and marital status.

The women's points of view about the barriers preventing the female teachers working in elementary schools from becoming school principals due to barriers resulting from family was (Mean=3,5018), and the men's points of view about the barriers preventing the female teachers working in elementary schools from becoming school principals due to barriers resulting from family were (Mean=2,8695). Consequently, there is a meaningful difference among the views of male and female teachers ($t:-5,82;p<.000$). The women's point of view about the barriers preventing the female teachers working in elementary schools from becoming school principals due to sex stereotypes was (Mean=2,9574); and the men's points of view about the barriers preventing the female teachers working in elementary schools from becoming school principals due to sex stereotypes were (Mean=2,6136). Consequently, there is a meaningful difference among the views of male and female teachers ($t:-3,00;p<.001$). The source of difference is the women's points of view in both of the dimensions. Compared with men, more women think that barriers resulting from family and societal sex stereotypes cause barriers to female teachers' career development.

Table 2*T-test Related with the Views of Participants in Terms of the Sex Variable*

	Sex	N	Mean	S	t	p
Resulting from family	Female	141	3,5018	,84025	5,82	,000**
	Male	91	2,8695	,75160		
School and environment	Female	141	2,4058	,84651	,38	,699
	Male	91	2,3663	,60303		
Working hour, Age and Marital Status	Female	141	2,8574	,78588	1,14	,254
	Male	91	2,7407	,71492		
Societal Sex	Female	141	2,9574	,95772	3,00	,001**
	Male	91	2,6136	,65588		
The women's opinion about administration	Female	141	2,5805	,78492	-1,37	,171
	Male	91	2,7206	,71426		
Economic Reasons	Female	141	2,7128	1,08436	,068	,946
	Male	91	2,7033	,93683		

**p<0.01

The third sub-aim of the research was expressed as: "According to the education variable, is there a meaningful difference among the barriers preventing the female teachers working in elementary schools from becoming school principals according to the views of teachers and school principals?" In Table 3, according to the level of education, the variance analysis and the results related to the third sub-aim are presented. There is a meaningful difference ($F:3,813$; $df:(3-228)$; $p<.011$) between the subjects' perception about the dimensions of working hours, age, and marital status in terms of the variable of education level. To determine the source of the difference, the Tukey HSD test was applied and it was found that the subjects who had a master's degree created the source of the difference. There aren't any meaningful differences among the subjects' perceptions in relation to other dimensions.

Table 3

Values of n, Mean, Standart Deviation and One Way Variance Analysis Related with the Views of Participants in Terms of the Variable of Education Level

Variables	Groups	N	Mean	sd	df	Mean Sq	F	p
Resulting from family	Associate degree	38	3,12	,97				
	Bachelor's degree	174	3,26	,85	3	,408	,545	
	Master's degree	12	3,44	,54	228	,748		,652
	Other graduates	8	3,35	,77	231			
	Total	232	3,25	,86				
School and environment	Associate degree	38	2,33	,86				
	Bachelor's degree	174	2,38	,73	3	,613		
	Master's degree	12	2,75	,74	228	,576	1,065	,365
	Other graduates	8	2,30	,65	231			
	Total	232	2,39	,75				
Working hour, Age and Marital Status	Associate degree	38	2,55	,88				
	Bachelor's degree	174	2,82	,71	3	2,122		
	Master's degree	12	3,37	,62	228	,556	3,813	,011*
	Other graduates	8	2,86	,84	231			
	Total	232	2,81	,75				
Societal Sex	Associate degree	38	2,75	1,03				
	Bachelor's degree	174	2,79	,81	3	,929		
	Master's degree	12	3,19	,88	228	,749	1,239	,296
	Other graduates	8	3,14	,98	231			
	Total	232	2,82	,86				
The women's opinion about administration	Associate degree	38	2,48	,79				
	Bachelor's degree	174	2,65	,74	3	,384		
	Master's degree	12	2,76	,85	228	,579	,663	,576
	Other graduates	8	2,64	,77	231			
	Total	232	2,63	,75				
Economic Reasons	Associate degree	38	2,56	1,14				
	Bachelor's degree	174	2,68	,98	3	2,084		
	Master's degree	12	3,37	1,06	228	1,041	2,002	,114
	Other graduates	8	2,812	1,09	231			
	Total	232	2,70	1,02				

Source of Difference (Tukey HSD), * $p < 0.05$

The fourth sub-aim of the research was expressed as: "According to the school principals' and teachers' views, do the barriers preventing the female teachers working in elementary schools from becoming school principals result in a meaningful difference in terms of the variable of the spouse's job?" In Table 4, according to the variable of the spouse's job, the variance analysis and the results that made for the meaningful difference between the subjects' values of mean and standard deviation and their views about the female teachers' career barriers are presented. There is a meaningful difference between the subjects' perception about the dimensions of societal sex stereotypes ($F:2,188;df:6-225;p<.045$), the women's opinion about making a career ($F:2,976;df:6-225;p<.008$), and the economic reasons ($F:2,457;df:6-225;p<.025$) in terms of variable of spouse's job. To determine the

source of the difference, the Tukey HSD test was applied. According to this result, the housewife, nurse, and other jobs constitute the source of the difference in societal sex stereotypes. The variable of civil servant and other jobs constitutes the source of difference in women's opinions about administration. While the nurse, engineer, and other jobs constitute the source of difference related to the economic reasons; it was noticed that there was not a meaningful difference between the subjects' perceptions related to other dimensions.

Table 4

Values of n, Mean, Standart Deviation and One Way Variance Analysis Related With the Views of Participants in Terms of the Variable of Spouse's Job

Variables	Groups	N	Mean	sd	df	Mean Sq	F	p
Resulting from family	Teacher	121	3,3626	,86141				
	Retired	9	3,3056	,83411				
	Civil servant	18	3,1736	,91669	6	,870	1,176	,320
	Housewife	27	2,9815	,67377	225	,740		
	Nurse	14	2,9107	1,23006	231			
	Engineer	10	3,2500	,50690				
	Other jobs	33	3,2538	,86404				
	Total	232	3,2538	,86238				
School and environment	Teacher	121	2,5041	,78764				
	Retired	9	2,4815	,66898				
	Civil servant	18	2,3704	,81293	6	,796	1,396	,217
	Housewife	27	2,2716	,60728	225	,570		
	Nurse	14	2,3492	,84716	231			
	Engineer	10	2,2667	,63764				
	Other jobs	33	2,1111	,71847				
	Total	232	2,3903	,75917				
Working hour, Age and Marital Status	Teacher	121	2,9446	,76681				
	Retired	9	2,5111	,53723				
	Civil servant	18	2,7000	,79261	6	1,105	1,964	,072
	Housewife	27	2,7741	,70580	225	,563		
	Nurse	14	2,4286	,89049	231			
	Engineer	10	2,9400	,53166				
	Other jobs	33	2,6212	,73389				
	Total	232	2,8116	,75941				
Societal Sex	Teacher	121	2,9986	,83707				
	Retired	9	2,8519	,68943				
	Civil servant	18	2,7315	,94823	6	1,596	2,188	,045*
	Housewife	27	2,5370	,62589	225	,729		
	Nurse	14	2,4524	1,01785	231			
	Engineer	10	2,8333	,68041				
	Other jobs	33	2,6061	1,01706				
	Total	232	2,8226	,86707				

The women's opinion about administration	Teacher	121	2,7769	,73376				
	Retired	9	2,4603	,69904				
	Civil servant	18	2,3968	,67817	6	1,633	2,976	,008**
	Housewife	27	2,8201	,82883	225	,549		
	Nurse	14	2,4796	,96664	231			
	Engineer	10	2,4429	,54896				
	Other jobs	33	2,2684	,67078				
	Total	232	2,6355	,75951				
Economic Reasons	Teacher	121	2,8967	,98935				
	Retired	9	2,7778	1,09291				
	Civil servant	18	2,8056	,94151	6	2,497		
	Housewife	27	2,5370	,93978	225	1,016	2,457	,025*
	Nurse	14	2,2857	,97496	231			
	Engineer	10	1,9500	1,18907				
	Other jobs	33	2,5000	1,09687				
	Total	232	2,7091	1,02693				

* Source of Difference (Tukey HSD), * p<0.05 **p<0.01

Conclusions and Recommendations

In the research that was conducted, the barriers preventing the female teachers from becoming school principals were studied. First of all, according to the position variable, a meaningful difference for the dimension of societal sex stereotypes was detected among the views of teachers and school principals. This finding can be explained by the fact that the number of male principals in the sample was greater than that of female school principals: 26 of the 30 school principals were men.

According to the sex variable, there is a meaningful difference among the views of the men and women in terms of the dimensions of the barriers resulting from the family and societal sex stereotypes. The barriers resulting from family can be explained by the fact that a women's priorities in today's societies are thought to be motherhood and being a housewife and a good spouse (according to men's points of view); while the activities of career development are regarded as being men's work (Ayan, 2000; Usluer, 2000). In the process of career development and appointment to the senior positions, it is necessary to work full time. It is very difficult for women to work full time while ignoring their responsibilities for their family (Ayan, 2000,s.41). Long working hours, the necessity of working on tasks which weren't finished at the office, taking care of the children (which is conceived as one of the most important duties for women), and the necessity of doing housework cause women to have difficulties in their career advancement. This situation causes women to perceive teaching as a traditionally female job which will enable them to carry out these duties (Altınışık, 1998; Usluer, 2000). Therefore, while the job of teaching (which includes shorter working hours and long holidays) is viewed as being a female job by society; the process of career development, which inhibits women from doing the traditional duties indicated above, isn't supported by the same society (Wilson, 2002). Consequently, it can be concluded that women don't want to enter the process of career development under the influence of men in superior positions who outnumber them and

the influence of a patriarchal culture which is still protected by men (Procter & Maureen, 1999, p. 153).

According to the results of the research, it can be understood that societal sex stereotypes, to a great extent, cause barriers to women becoming school principals. The societal sex stereotypes, which dominated during the Cold War, obviously have had an influence on contemporary societies. The stereotypes which dictate that men should be teachers and school principals have a significant influence on families. Families prefer male school principals, or female school principals who behave like men, because the parents consider that a serious atmosphere is created by authoritarian male school principals, or by female school principals who are authoritarian like their male counterparts (Smulyan, 2000). The parents of students generally think that female school principals cannot discipline male students; and therefore they are concerned about discipline problems (Coleman, 2005). The reasons indicated above cause society to view male school principals as being superior, and these reasons also cause female school principals to behave in the same way as men (Conner & Sharp, 1992). Moreover, factors like women's acceptance of the idea that men manage the family, which has been a norm for centuries; cause the job of administration to be regarded as a man's job (Tan, 1996). There are research studies which support the findings and the results of this study in Turkey and in other countries. For example, in Hungary, women constitute 85 percent of the total number of teachers; whereas female school principals constitute 33 percent of the total number of school principals. In the Netherlands, 76 percent of the teachers are female and 13 percent of the school principals are female (Athanasoula-Reppa, A. & Koutouzis, M, 2002). In the USA, women comprise 56% of those employed in education. However, the fact that a high percentage of educators are women doesn't support the view that they have the same rights as men have, or that they can join the labour force as easily as men can. This is because the kind of numerical data given above confirms the existence of sex discrimination against women when compared with the numerical data related from other areas of employment.

The numerical data shows that women in the USA occupy 4% of the senior positions in the most important companies; they constitute 13% of the senators, they constitute 2% of high ranking military personnel; and they constitute 30% of lawyers in the legal field (Carli & Eagly, 2001, pp. 630-631). This data confirms the existence of societal sex stereotypes in contemporary societies; such as the ones which indicate that women can only become a teacher, that the teaching profession suits women more than any other job, that women can't fulfill the duties required by senior positions, and that they can't demonstrate the attributes of leadership which are necessary for these positions. In particular, the women who want to reach the top administrative positions face a glass ceiling and, consequently, they are represented in limited numbers in the administrative positions which require a great deal of responsibility (Arıkan, 2003). As was explained previously, the number of female school principals constitutes 3 percent of the total number of school principals. In a research study carried out in 1996, Tan found out that men constitute the majority of the high ranking administrators in the General Directorship of Preschool Education and the General Directorship of Girl Technical Education; despite the fact that these institutions were generally dealing with women (1996). Furthermore, the fact that teaching is perceived as a woman's job, while the job of

educational administration is regarded as a man's job, is a stereotype which has been in existence for a long time. Therefore, male participants may have different points of view than those of female participants. Consequently, the prejudice and discrimination which have been created against women are effectively reinforced by society and continue to exist (Charles & Davies, 2002; Kerr, Miller & Reid, 2000).

Moreover, according to the variable of school principals and teachers' education level, there is a meaningful difference between their views about the dimensions of working hours, age, and marital status. This difference results from the views of the subjects who have a master's degree. The subjects with a master's degree state that they agree with the dimensions of working hours, age, and marital status because of their experience of the problems during the process of their career development.

Lastly, according to the variable of a spouse's job, there is a meaningful difference between the views of teachers and school principals about the dimensions of societal sex stereotypes, and women's opinions about administration and economic problems. According to this finding, the source of the differences in societal sex stereotypes are those whose spouses are bank employees and nurses. The barriers opposing women during the period of career development in the banking sector and health sector resemble the barriers that women experience in the educational field. The women in these sectors have long working hours and they don't have enough opportunities to be promoted (Charles & Davies, 2000, pp. 547-548; Reed & Fischer, 2001, p.139; Lane, 2000, p. 269-272).

Recommendations

According to the sex variable, the point of view of teachers is different from those of school principals in terms of the barriers resulting from families and societal sex stereotypes. Compared with male teachers, more female teachers state that they experience barriers in terms of career advancement towards becoming school principals. In particular, male teachers and school principals should understand the emotions and opinions of female teachers regarding familial subjects and societal sex stereotypes, and they should help the female teachers to break these stereotypes.

The authorities in the highest positions of the Ministry of National Education should contribute to the elimination of the obstacles which prevent women from becoming school principals, and they should support women in becoming school principals. For example, the laws should be reorganized so that an equal number of male and female teachers can be school principals. Moreover, it should also be realized that female teachers and other individuals in an organization have an important responsibility to help female teachers to overcome the barriers to becoming school principals. The sample for this research was limited to the centre of Mersin Province. This study, which has been carried out in a fairly limited area, can be done using a larger sample. Also, the issue of whether the career barriers that the female teachers have experienced negatively influence their work performance can be studied.

References

- Altınşık, S. (1998). *Kadın öğretmenlerin okul müdürü olmasının engelleri* [The barriers to female teachers' becoming school principals]. Unpublished doctoral dissertation, Hacettepe University, Institute of Social Sciences, Ankara.
- Arıkan, S. (2003). Leadership behaviours of female administrators and an application in banking sector. *I.B.B.F Journal of Gazi University*, 1, 1-19.
- Athanassoula-Reppa, A. & Koutouzis, M. (2002). Women in managerial positions in Grek education: Evidence of inequality. In P. V. Glass (Ed.), *Education Policy Analysis Archives*, 10, 11.
- Ayan, F. (2000). *Bayan öğretmenlerin yöneticilikleri ve öğretmen görüşleri* [Female teachers' administration and the opinion of teachers]. Unpublished master's thesis, Çanakkale Onsekiz Mart University, Institute of Social Sciences, Çanakkale.
- Carli, L.L. & Eagly, A. H. (2001) Gender, hierarchy, and leadership: an introduction. *Journal of Social Issues*, 57, 629-636.
- Charles, N. & Davies, C. A. (2000) Cultural stereotypes and gendering of senior management. *The editorial Board of the Sociological Review*, 544-567.
- Coleman, M. (2005). Gender and secondary school leadership. *Gender Issues in Leadership*, 33, 3-20. Retrieved May 16 2006 from <http://search.epnet.com/login.aspx?direct=true&db=aph&jid=CER>
- Comer, N. L. & Sharp, W. L.(1992). *Restructuring schools: will there be a place for women?*. Retrieved May 3 2006 from <http://search.epnet.com>.
- Çelikten, M. (2004). Okul müdürü koltuğundaki kadınlar: Kayseri İli [The women at principals' chair: Kayseri Province]. *E.Ü. Eğitim Fakültesi Sosyal Bilimler Enstitüsü Dergisi*, 2, 91-118.
- Çelikten, M. (2005). The women at principals' chair in Turkey. *The Journal of American Academy of Business*, 6, 85-94.
- Evetts, J. (2000). Analysing change in women's careers: culture, structure and action dimensions. *Gender, Work and Organization*, 7, 57-67.
- Hudson, J. & Rea, D. (1998). Teacher's perceptions of women in the principalship. A current perspective. *Advancing Women In Leadership*. Retrieved June 1 2006 from <http://www.advancingwomen.com/awl/summer98/HUD.html>.
- ILO. (1997) *Women's progress in workforce improving worldwide, but occupational segregation still life-"glass ceiling" separates women from top jobs*. Retrieved May 15 2006 from <http://www.ilo.org/public/english/bureau/inf/pr/1997/35.htm>.
- Kerr, B., Miller, W. & Reid, M. (2002). Sex-based occupational segregation in U.S. state bureaucracies. *Public Administration Review*, 62, 412-423.
- Kadının Statüsü ve Sorunları Genel Müdürlüğü. ([KSSGM], 2000). *Kadın istihdamı için yeni perspektifler ve kadın işgücüne muhtemel talep* [New perspectives for female employment ve the possible demand for female workforce.]. Ankara: Author.

- Lane, N. (2000). The low status of female part-time NHS nurses: a bed-pan ceiling. *Gender, Work and Organization*, 7, 269-281.
- MEB. (2006). Mersin Milli Eğitim Müdürlüğü İstatistik Birimi.
- Moreau, M.P., Osgood, J. & Halsall, A. (2007) Making sense of the glass ceiling in schools: an exploration of women teachers' discourses. *Gender and Education*, 19 (2), 237-253
- Merle, A. (1999). *Women's journeys to educational administration positions*. Retrieved May 4 2006 from <http://www.ssta.sk.ca/research/leadership/00-08.htm>.
- Oplatka, I. (2001). I changed my management style: the cross-gender transition of women headteachers in mid-career. *School Leadership & Management*, 2, 219-233.
- Otaran, N., Sayin, A., Güven, F., Gürkaynak, İ., & Atakul, S. (2003). *A gender review in education, Turkey 2003. Education process (teachers & principals)*. Retrieved May 13 2006 from <http://www.unicef.org/turkey/gr/ge21g.html>.
- Özbilgin, F. M. & Woodward, D. (2004). Belonging and otherness: sex equality in banking in Turkey and Britain. *Gender, Work and Organization*, 11, 668-688.
- Procter, I. & Maureen, P. (1999). Work orientation and women's work: a critique of Hakim's theory of the heterogeneity of women. *Gender, Work and Organization*, 6, 152-161.
- Reed, V. & Buddeberg Fischer, B. (2001). Career obstacles for women in medicine: an overview. *Medical Education*, 35, 139-147.
- Rimmer, R. & Rimmer, S. (1997). Employment breaks, pay and career development among Australian women. *Gender, Work and Organization*, 4, 202-217.
- Smulyan, L. (2000) Feminist cases of nonfeminist subjects: case studies of women principals. *Qualitative Studies in Education*, 13, 589-609.
- Tan, M. (1996). Eğitim yönetiminde kadın azınlık. *Amme İdaresi Dergisi*, 4, 35-36.
- Thompson, A. (2003). Special series on girls and women in education caring in context: four feminist theories on gender and education. *The Ontario Institute for Studies in Education. Curriculum Inquiry*, 33, 1. Oxford: Blackwell Publishers.
- Thompson, B. (2007). Working beyond the glass ceiling: women managers in initial teacher training in England. *Gender and Education*, 19, 339-352.
- Usluer, L. (2000). *Kadın öğretmenlerin yönetici konumlara yükseltilmeme nedenleri konusundaki öğretmen ve yönetici görüşlerinin değerlendirilmesi* [The evaluation of the views of teachers and administrators about the reasons for preventing female teachers from being promoted to administrators]. Unpublished doctoral dissertation, Ankara University, Ankara.
- Wilson, F. M. (2002). Management and the professions: how cracked is that glass ceiling?. *Public Money & Management*, 5, 15-20.
- Zafarullah, H. (2000) Through the brick wall, and the glass ceiling: women in civil service in Bangladesh. *Gender, Work and Organization*, 7, 97-209.

Kadın Öğretmenlerin Okul Müdürü Olmaları Önündeki Engeller (Mersin İli Örneği-Türkiye)

(Özet)

Problem Durumu: Sanayi devriminden sonra kadınlar bazı haklar elde etmeye başladılar ve bu hak elde etme süreci yirminci yüzyılda da devam etti. Örneğin, kadınlar yirminci yüzyılda birçok ülkede oy verme hakkı ve mülk edinme haklarını elde ettiler. Birçok ülkenin anayasası kadın erkek eşitliğini kabul etti. Bütün bu olumlu gelişmelere rağmen kadınlar erkeklerle gerçek anlamda eşit olamadılar. Kadın ve erkek arasındaki eşitsizliğin devam ettiği alanlardan birisi de eğitim alanıdır. Eğitim kurumlarında kadınlar erkeklerle kıyaslandığında okul müdürlüğü gibi mevkilere ve üst düzey mesleklere terfi etme şansına daha az sahiptir. Bundan dolayı yönetici pozisyonunda olan kadınların sayısı işgücüne katılan kadınlara oranla çok azdır. Eğitim alanında Türkiye’de de kadınlara benzer bir eşitsizlik söz konusudur. Dünyada ve Türkiye’de kadınların kariyerlerinde ilerlemelerini engelleyen nedenler benzerlik göstermektedir. Birçok araştırmaya göre bu engeller şunlardır:

- a- Cam tavan
- b- Toplumsal cinsiyet kalıp yargılar kaynaklı;
- c- Aile kaynaklı;
- d- Kadınların yöneticiliğe bakış açıları;
- e- Eğitim, çalışma saati, yaş ve medeni durumdan kaynaklı;
- f- Okul ve çevre kaynaklı kariyer engelleri.

Bu çalışma ilköğretimde görevli öğretmen ve müdür görüşlerine göre ilköğretimde çalışan kadın öğretmenlerin okul müdürü olma engellerini incelemektedir.

Araştırmanın Amacı: Araştırmanın amacı resmi ilköğretim okullarında çalışan öğretmenlerin ve müdürlerin görüşlerine göre ilköğretimde çalışan kadın öğretmenlerin okul müdürü olması önündeki engelleri belirlemektir. Ayrıca bu araştırmada kadınların yaşadıkları okul müdürü olma engelleri konusunda görüşlerini belirten katılımcıların görüşleri arasında konum, cinsiyet, eğitim durumu, eşin mesleği, değişkenleri açısından anlamlı bir farklılık olup olmadığının belirlenmesi hedeflenmektedir.

Araştırmanın Yöntemi: Betimsel nitelikte olan bu araştırmada genel tarama modeli kullanılmış, örneklem olarak orantısız eleman örnekleme kullanılmıştır. Araştırmanın çalışma evrenini, 2006-2007 öğretim yılında Mersin ilindeki ilköğretim okullarında görevli toplam 150 müdür ve 4200 öğretmen oluşturmaktadır. Toplam 150 müdürden 41’ine ulaşılmış ancak müdürlerin bir kısmı ölçeği doldurmadığı için bir kısmı da eksik ve hatalı doldurdukları için 11 ölçek araştırma kapsamı dışında tutulmuştur. Toplam 4200 öğretmenden 225’ine ulaşılmış ancak öğretmenlerin 23 tanesi ölçek maddesini eksik ve hatalı doldurmaları nedeniyle araştırma kapsamı dışında tutulmuştur. Belirtilen ildeki okullardan toplam 30 yönetici ve 202 öğretmen yansız olarak seçilmiş ve toplam 232 müdür ve öğretmen araştırmanın örneklemini oluşturmuştur.

Araştırmanın Bulguları: Bulgulara göre konum değişkeni toplumsal cinsiyet kalıp yargıları boyutunda anlamlı bir farklılığa neden olmuştur. Konum değişkeni açısından müdür ve öğretmenlerin görüşleri arasında anlamlı bir farklılık vardır. Farklılığa ise öğretmenlerin algıları neden olmuştur. Yine cinsiyet değişkeni

ailevi nedenler ve toplumsal cinsiyet kalıp yargılar boyutunda anlamlı bir farklılığa neden olmuştur. Kadınlar, erkeklere göre, kadın öğretmenlerin okul müdürü olma önünde ailevi nedenlerin ve toplumsal cinsiyet kalıp yargıların daha çok engel olduğunu düşünmektedirler.

Müdür ve öğretmenlerin eğitim durumu değişkenine göre "çalışma saati, yaş, medeni durum" boyutuna yönelik algıları arasında (F:3,813;sd:(3-228);p<.011), anlamlı bir farklılık vardır. Farkın kaynağını yüksek lisans mezunu öğretmenler oluşturmuştur. Bulgulara göre müdürlerin ve öğretmenlerin eşin mesleği değişkenine göre "toplumsal cinsiyet kalıp yargılar (F:1,596;sd:6-225);p<.045)", "kadınların kariyer yapmaya bakış açıları (F:1,633;sd:6-225);p<.008)," ve "ekonomik nedenler (F:2,457;sd:6-225);p<.025)," boyutuna yönelik algıları arasında anlamlı bir farklılık vardır. Buna göre toplumsal cinsiyet kalıp yargılarda farkın kaynağı ev hanımı, hemşire ve diğer meslek grupları oluşturmaktadır. Kadınların kariyere bakış açılarında farkın kaynağını memur ve diğer meslek grupları değişkeni oluşturmaktadır. Ekonomik nedenlere ilişkin olarak farkın kaynağını ise hemşire, mühendis ve diğer meslek grupları oluştururken diğer boyutlara ilişkin deneklerin algıları arasında anlamlı bir farklılık olmadığı görülmüştür.

Araştırmanın Sonuçları ve Önerileri: Araştırma sonuçlarına göre öğretmenler müdürlere göre toplumsal cinsiyet kalıp yargıların kadınların müdür olmaları önünde daha fazla engel teşkil ettiğini belirtmişlerdir. Cinsiyet değişkenine göre, kadın öğretmen ve müdürler, erkek öğretmen ve müdürlere oranla "ailevi nedenler ve toplumsal cinsiyet kalıp yargılar" boyutlarında kadınların okul müdürü olma önünde daha fazla engel oluşturduğunu belirtilmiştir.

Müdür ve öğretmenlerin "eğitim durumu" değişkenine göre "çalışma saati, yaş, medeni durum" boyutuna yönelik görüşleri arasında anlamlı bir farklılık vardır. Bu fark yüksek lisans mezunu öğretmenler ve müdürlerin görüşlerinden kaynaklanmaktadır. Çünkü yüksek lisans mezunu öğretmen ve müdürlerin kariyerlerini geliştirme sürecinde bu tür problemleri bizzat yaşamış olduklarından dolayı diğer eğitim düzeyindeki katılımcılardan farklı düşünülmüş olabilirler.

Müdür ve öğretmenlerin eşin mesleği değişkenine göre, toplumsal cinsiyet kalıp yargılar, kadınların kariyer yapmaya bakış açıları ve ekonomik nedenler boyutuna yönelik görüşleri arasında anlamlı bir farklılık vardır. Bu bulguya göre toplumsal cinsiyet kalıp yargılarda farkın kaynağını, eşi bankacı ve hemşire olanların oluşturduğu görülmektedir. Banka ve sağlık sektörlerinde kadınların kariyer gelişim engelleri, eğitim alanında kadınların yaşadıkları engellere benzemektedir.

Bütün bu sonuçlar göz önüne alındığında cinsiyet değişkenine göre öğretmen ve müdürlerin görüşleri ailevi konularda ve toplumsal cinsiyet kalıp yargılar konularında değişmektedir. Kadın müdür ve öğretmenler erkelere göre daha fazla engelle karşılaştıklarını belirtmektedirler. Bu nedenle erkek müdür ve öğretmenlerin kadın öğretmenlerin duygu ve düşüncelerini anlamları bu kalıp yargıların kırılmasında kadın öğretmenlere destek olmaları sağlanmalıdır.

Milli Eğitimin üst kademesinde görev alan yetkililerin kadınların okul müdürü olmaları önündeki yasal olmayan engellerin kaldırılmasına katkı sağlaması ve kadınların okul müdürü olmalarını özellikle desteklemeleri gerekir.

Anahtar Sözcükler: Kariyer engelleri, okul müdürü, öğretmen, cinsiyet kalıp yargılar

Reliability and Validity Study of Turkish Form of the Psychological Distress Scale

Hülya M. Ünal Karagüven*

Suggested Citation:

Karagüven, M. H. (2009). Reliability and validity study of Turkish form of the psychological distress scale. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 36, 179-192.

Abstract

Statement of Problem: Traumatic events are conceptualized as environmental stressors and affect large numbers of people throughout the world. Psychological problems may arise after being exposed to a stressful event. It is important to have reliable instruments with which to address traumatic events' long term consequences.

Purpose of the Study: The purpose of this study is to adapt the Psychological Distress Scale (PDS) into Turkish by examining the results of reliability and validity analyses results of Turkish form of the PDS.

Method: This study involved 1144 adults and young people. The self-report PDS was translated into Turkish and named as Psikolojik Yoğun Stres Ölçeği (PYSÖ). Data was collected within two months of the Marmara Earthquake, which occurred in the Marmara region of Turkey in August 1999. Outcome variables also included the measures of state-trait anxiety and Impact of Event. Correlation coefficient, T-test, analyses of variance, Factor analysis, Cronbach's alpha, item-total correlation, and standard error of measurement techniques were used to analyze the data.

Findings and Results: In the preliminary analyses, alpha was found to be .79 for the original English PDS forms. Following the translation and back-translation process, for transliterational equivalence adequate correlation values obtained between English and Turkish forms. For criterion-related validity, to examine the construct validity of the PYSÖ, statistically significant correlation values obtained between PYSÖ and Turkish form of State-Trait Anxiety Inventory-STAI. Results suggested that the constructs were relevant to each other. Statistically significant relationships with other similar measures provided evidence for the concurrent validity. For predictive validity, hypotheses testing method was used. Results showed that PYSÖ mean scores were significantly different for gender, experiencing area, removing time from disaster area, having lost a close one, and being a mother. Analysis of variance results showed that age,

* Asst. Prof. Dr., Marmara University Faculty of Technical Education, mhulya@marmara.edu.tr

education level, job, and area of residence were also significant factors in the mean scores. According to factor analysis results, PYSÖ had two dimensions but could also be used as a one-dimensional general scale. In order to test reliability, Cronbach's alpha and item-total correlation coefficients techniques were used to assess the internal consistency. The Cronbach's alpha score for the PYSÖ was .83. This was also very similar with Turkish form of the Impact of Events Scale's alpha scores which were obtained from the same groups. These results showed that translated Turkish forms of scale were easily understandable by Turkish groups. Item-total correlation was ranging from .79 to .82 for total group and subgroups. The standard error of measurement technique was also used for reliability analyses and was found to be 2 points at .01 level.

Recommendations: PYSÖ can be used to evaluate psychological distress in psychometric assessment process of related studies. Further study is recommended in order to assess the reliability and validity of PYSÖ with different populations and in different stressful events.

Keywords: Psychological distress scale, validity, reliability

Reliability and validity studies are topics of major interest for education researches. Reliable and valid Turkish instruments provide opportunities in order to collect reliable and valid data for researchers. Because of the limited equivalent Turkish measurements, researchers need to work on validation of measurements which are originally developed in different languages. Especially in traumatic event exposure, studies there have had very limited measures to investigate psychological effects of traumatic events. Turkey suffered a high magnitude earthquake in August 1999 and thousands of people lost their lives in the Marmara region. Numbers of people were affected physiologically and psychologically from this disaster. These kinds of events impact large numbers of people not only in Turkey, but also throughout the world. Only in the United States, it has been estimated that as many as two million adults suffer injury or property damage each year due to such events (Solomon, 1989). According to 2008 World Health Organization, in recent years statistics there have increased in accident and injury in high, middle, and low income countries (World Health Organization [WHO], 2008). Studies have indicated that these kinds of events have negative impacts on mental health (Freedy & Hobfoll, 1995).

Hobfoll (1988, 1989, 1998) has suggested in Conservation of Resource (COR) theory that extreme events result in widespread loss of resources which, in turn, may lead to stress. As a result of resource loss, stress may occur. These kinds of events that happen to a community are termed traumatic or stressful events and conceptualized as environmental stressors (Hobfoll, 1995, 2001).

Psychological problems may arise after a stressful event exposure. Initial reactions to an unexpected traumatic event may include surprise, helplessness, shock, anger, and confusion, although many of these reactions are resolved in most survivors within a few months (Freedy, Kilpatrick & Resnick, 1992). Post-Traumatic

Stress Disorder (PTSD) symptoms, higher distress levels, anxiety disorders, somatization, and depression have been generally reported as long term mental health consequences from traumatic events (Green, 1990; Solomon, 1989). People who experience high levels of trauma are likely to experience psychological distress. Psychological distress is a state of not only depression but also anxiety. Depressed individuals tend to be anxious while anxious individuals tend to be depressed in a psychological, distressed situation (Mirowsky & Ross, 1989). Different methods are available to assess mental health problems after a traumatic event exposure. Questionnaires and scales have very useful roles in the psychometric assessment process. In order to create a scale, based on individual claims, modified items from these two types of psychological discomfort situations are listed as mood and malaise. There have been some well-known, structured, and standardized instruments originally developed in this area.

One of the very widely used devices for assessing post-traumatic stress is 17 Post-Traumatic Stress Disorder (17 PTSD). PTSD is assessed utilizing the frequency of experiencing 17 PTSD symptoms derived from the Diagnostic and Statistical Manual (American Psychiatric Association [APA], 1987). The Clinician-Administered PTSD Scale (CAPS) (Blake et al., 1990) is specifically designed to yield both continuous (i.e., severity) and dichotomous (i.e., diagnostic) information exclusively for PTSD. Symptom Checklist-90 Revised (SCL-90R) (Derogotis, 1983) is a 90-item self-report inventory designed to assess general psychological distress. General levels of distress can also be measured using instruments such as the 28-item version of the General Health Questionnaire (Goldberg, 1972). The Impact of Events Scale (IES) (Horowitz, Wilner & Alvarez, 1979) is probably the most widely used self-report measure in the field of traumatic stress. It is a 15-item scale assessing subjective distress after a stressful life event. It has two subscales of Intrusion and Avoidance. This Scale was translated into Turkish as "Olayın Etkisi Ölçeği" (OEÖ) and used in several studies after the Marmara Earthquake of 17 August 1999 (Karagüven & Erkin, 2000; Karagüven & Erkin, 2001).

More recent examples of these types of instruments include a 23-item Psychological Distress Manifestations Measure Scale developed by Paulin, Lemoine, Poirer and Lambert (2000) and a 15-item psychological distress scale developed by Patterson (2005).

Despite all of these types of measures, it is still necessary to have a simple and short reliable instrument to determine the consequences of a stressful event exposure. Psychological Distress Scale (PDS) was developed by Freedy, Kilpatrick, Resnick, and Saunders (1994). This 9-item scale has the advantage of being easy to administer and inexpensive to use. An easily scored instrument can be administered alone or together with other measures. The aim of the present study was to examine the results of reliability and validity studies for the Turkish form of PDS in order to introduce a valid and reliable equivalent Turkish instrument for Turkish researchers.

Method

Sample and Procedure

Two separate sample groups were studied: undergraduate sample and community sample. For the transliteral equivalence study, the undergraduate sample consisted of 100 bilingual university students from English Language Departments of the following two different universities: Boğaziçi and Marmara in Istanbul. The scale was translated into Turkish by bilingual lecturers in two different universities and expert opinion was obtained for the Turkish items. Turkish form of PDS was entitled "Psikolojik Yoğun Stres Ölçeği" (PYSÖ). Undergraduate participants were administered two different test batteries in group settings on two occasions respectively within 2 weeks: first, the English batteries of PDS; and second, translated Turkish batteries of PDS. The community group consisted of 1403 residents from tent cities in earthquake area and Anatolian side of Istanbul. The sample size was different due to the missing values. The number of subjects varied from 1403 to 1144. Correlation coefficient method was used in this study. Data were collected within two months after the Marmara earthquake. Of the data, 70% was gathered from Istanbul and 30% from the disaster area (tent cities). The sample was determined with two big tent cities in disaster area and Anatolian side in Istanbul. Participation was voluntary and individuals were recruited without regard to age or gender. In the community sample, average age was 24 ($X=24.11$, $S=9.09$, $min.=15$, $max.=55$). Of the participants, 60% were male. Of the female participants, 24% were mothers. Across the sample, 1% was illiterate, 5% was primary school graduated, 14% was secondary school graduated, 66% was high school graduated, 10% was university graduated, and 4% had a higher degree of the sample (see table1). All participants were informed about consent.

Measures

Questionnaire. Demographic and specific characteristics were of participants determined by a 19-item questionnaire.

Psychological distress scale. Psychological Distress Scale (PDS) (Freedy et. al., 1994) was a 9-item symptom checklist, originally developed in English. Items assessed intrapsychic and interpersonal complaints such as intrusive and frightening thoughts, irritability, depressed effect, sleep difficulty, hopelessness, tachycardia, temper outburst, worrying, and feeling critical of others. Participants reported on complaint intensity during the past weeks. The average item response was used as an index of current psychological distress (0 = not at all, 1 = a little bit, 2 = quite a bit). The original form has high alpha values. For a group of adult sample ($N=299$), Coefficient alpha was found as .82 (Freedy et al., 1994).

State-trait anxiety inventory. STAI was developed by Spielberger, Gorsuch, and Lusbene (1970). The Turkish form of STAI was studied by Öner and Le Compte (1985) and entitled as "Durumluk-Sürekli Kaygı Envanteri" (DSKE). This inventory is a well validated and widely-used Turkish measure to evaluate state and trait anxiety (Öner, 2006). The inventory consists of two forms, each being 20 items in length. First 20 items assess state anxiety and second 20 items assess trait anxiety.

Table 1**Demographic Variables of Sample**

Variables		N	%
Area	Istanbul	1015	72,3
	Earthquake Area	382	27,2
Missing		6	,4
Total		1403	100,0
Gender	Women	562	40,1
	Men	817	58,2
Missing		24	1,7
Total		1403	100,0
Age	18-28	1033	73,6
	29-38	149	10,6
	39 +	119	8,5
Missing		102	7,3
Total		1403	100,0
Job	Student in high school	198	14,1
	Student in Univ.	796	56,7
	Lecturer in Univ.	73	5,2
	Teacher in high school	88	6,1
	Others	194	14
Missing		54	3,9
Total		1403	100,0
Edu. Level	Illiterate	10	,7
	Primary school	65	4,6
	Secondary school	196	14,0
	High school	892	63,6
	University	135	9,6
	Master and PhD	63	4,5
	Total	1361	97,0
Missing		42	3,0
Total		1403	100,0

Item remainder correlations were higher than .30: .34-.72 for trait anxiety and .42-.35 for state anxiety. Test retest reliability was also studied. Test retest reliability coefficients ranged between .76 and .86 for trait anxiety, and .26 and .68 for state anxiety (Öner & Le Compte, 1985, pp.11-12). Participants rated the degree to which they felt a given way over the previous week and in general. Ratings were made on a 4-point Likert scale, from 1=not at all to 4=very much so for the state form and the trait from 1=not at all to 4=very much so for the trait form. Both state and trait anxiety were used in the present study.

Impact of events scale. This scale was developed by Horowitz et al. (1979) and is a 15-item scale assessing subjective distress after a stressful life event. It has two subscales of Intrusion and Avoidance. This Scale was translated into Turkish as "Olayın Etkisi Ölçeği" (OEÖ) and used in several studies after the 17 August 1999 Marmara Earthquake (Karagüven & Erktin, 2000; Karagüven & Erktin, 2001). Respondents were asked to rate the items on a 4-point scale according to how often each had occurred in the past 7 days. The 4 points on the scale are as follows: 0 (not at all), 1 (rarely), 3 (sometimes), and 5 (often).

Results

At the beginning of the study, for the transliteral equivalence, the original English forms were administered to 100 bilingual university students. Alpha found .79 for this group. Secondly, PYSÖ forms were administered to the same group and the correlation values were .78 ($P \leq .01$) for total score. For the items, correlations were ranging from .64 ($P \leq .01$) to .76 ($P \leq .01$) as shown in Table 2. From these high and statistically significant correlation values it can be concluded that the Turkish and the English forms of the scale were significantly related and similar to each other.

Table 2

Correlation Coefficients Between Turkish and English Items of the PDS

Items	N=99
1.	.699**
2.	.702**
3.	.640**
4.	.668**
5.	.762**
6.	.644**
7.	.691**
8.	.715**
9.	.742**

** $P \leq .01$

Validity

For criterion-related validity, to examine the concurrent validity of the PYSÖ, the related construct of DSKE (Öner, 1985) was administered along with the PYSÖ. The correlation between PYSÖ and DSKE was statistically significant for both subscales: State Anxiety ($r = .33$, $P \leq .01$) and Trait Anxiety ($r = .40$, $P \leq .01$), suggesting that the constructs were relevant to each other as can be seen in Table 3. Significant relationship with other measures provided evidence for the concurrent validity of the scale.

Table 3

Correlation Coefficients Between Turkish Forms of PDS and STAI

	N	Trait Anxiety (r)	State Anxiety (r)
PDS	124 Earthquake area	.407**	.331**
	439 Istanbul	.335**	.323**
	564 Total	.352**	.309**

** $P \leq .01$

For distinctive validity, hypotheses testing method was used. T-test results showed that there were statistically significant differences for gender ($t=10.99$, $sd=4.19$, $P < .001$), between people who experienced the event in the most affected area, and those who experienced the event in a less affected area ($t=-6.34$, $sd=4.16$, $P < .001$), between people who left the disaster area immediately after the event and

those who did not ($t=-4.25$, $sd=4.28$, $P<.001$), between people who lost their close ones and those who did not ($t=-6.12$, $sd=4.29$, $P<.001$), and between women who were mothers and those who were not ($t=-7.11$, $sd=4$, $P<.001$). The analyses of variance results showed that age ($F_{2,1082}=17.732$, $P<.000$), education level ($F_{5,1129}=16,285$, $P<.000$), job ($F_{5,1112}=28.796$, $P<.000$), and residence area at present ($F_{9,1123}=18.118$, $P<.000$), were statistically significant.

For the construct validity, the explanatory factor analysis method was used. Principal component analysis conducted for extraction and Varimax rotation with Kaiser normalization was used to extract the factors from the items. According to the results of the preliminary factor rotated component matrix, factor loading of item 6 was (9,462E-02) less than .10 in the 2. factor. Therefore, item number 6 (tachycardia) was eliminated in the second factor analyses. Two factors, the eigenvalues of which are greater than 1.00, accounting 58.39% of the total variance were extracted. Results of factor analyses are shown in Table 4.

As seen in Table 4, items with factor loadings from .61 to .78 were loaded on one of the two factors. Factor 1 explained 32.7% of the variance and contained 5 items. This factor was entitled "Worry." Factor 2 explained 25.5% of the variance and consisting of 3 items. This factor was entitled "Emotionality." Because of the higher total variance (58.39%), the scale could also have only one general factor. Additionally, before rotated, Factor 1 showed high loadings (component matrix: .58-.71) even before item 6 (tachycardia) was eliminated (component matrix: .56-.72). Therefore, the scale could be used as a nine-item and one-factor general scale.

Table 4

Rotated Factor Loadings for the PYSÖ, Communalities (h^2), Numbers, Means and Standard Deviations

	Factor 1	Factor 2	h^2	N	M	Sd.
8. Worrying	.781	.137	.630	1186	2.20	.67
1. Intrusive and frightening thoughts	.756	.203	.614	1180	2.01	.72
5. Hopelessness	.697	.171	.515	1181	1.79	.73
3. Depressed effect	.660	.245	.496	1180	2.21	.73
4. Sleep difficulty	.611	.338	.488	1185	2.04	.80
7. Temper outburst	.177	.853	.759	1177	1.70	.75
2. Irritability	.253	.670	.791	1183	1.85	.71
9. Feeling critical of others	.217	.669	.495	1158	1.79	.78

Note: Factor loadings which are greater than .400 are in **bold**.

Reliability

For reliability studies of the Turkish version of the PDS, internal consistency coefficient and standard error of measurement techniques were used. Cronbach's (1951) alpha measures, means, standard deviations and item-total correlation coefficient values were calculated for internal consistency coefficient. Cronbach's alpha of the PYSÖ were .82 and .83 as shown in Table 5. This compared with alpha value of .78 for the English version. Alpha was .82 for the original English form in a

study with 229 adults (Freedy et al., 1994, s.264) Therefore, it can be said that the internal consistency reliability of the PYSÖ was found to be very similar to the original PDS. As can be seen in Table 6, the item-total correlation values ranged from .78 to .82 for total group and subgroups.

Table 5*Alpha Measures, Means, and Standard Deviations of PYSÖ*

Groups	Total Test (n=9)			
	N	Alpha	Mean	S.D.
Istanbul	805	.82	16.61	4.15
Earthquake Area	339	.82	18.39	4.41
Total	1144	.83	17.14	4.30

Table 6*Item-Total Correlations Coefficients of PYSÖ*

Items	Earthquake Area	Istanbul City	Total Scores
	N=339	N=805	N=1144
1	.78	.80	.80
2	.80	.80	.81
3	.79	.81	.81
4	.79	.80	.80
5	.80	.81	.81
6	.80	.82	.82
7	.80	.81	.81
8	.79	.80	.80
9	.81	.82	.82

Standard error of measurement technique was also used for reliability analyses. It was found as 2 points as seen in Table 7 for PYSÖ. Thus, the difference between an individual's score should be 2 points to be significant at .01 level. This means that a score can change approximately 2 points.

Table 7*Standard Error of Measurement Scores of PYSÖ*

Sample Groups	N	Standard Error Scores
Total	1144	2.05
Istanbul City	805	1.98
Earthquake Area	339	2.11

Conclusions and Recommendations

Traumatic events such as natural disasters, terrorism, and accidents impact large numbers of people, not only in Turkey but also in other countries. There are very few Turkish instruments to assess post traumatic stress disorder. The aim of this study

was to adapt the Psychological Distress Scale to Turkish. This can be useful for further research in this field. Firstly, the existing English forms were administered to a group of bilingual university students from English Language Departments. The alpha was .79 for this sample. This alpha value was very close to the original English form's alpha .82. (Freedly et al., 1994). This means that the English items were understandable for Turkish university students. Secondly, the scale was translated into Turkish and entitled as PYSÖ. Thirdly, for transliterational equivalence Turkish forms of the scale were given to the same groups of bilingual university students within two weeks respectively. High correlation values showed that the English and Turkish forms were very similar to each other. Fourthly, in order to examine the content related construct validity of the PYSÖ, the related and well-validated DSKE was administered along with the PDS. Significant relationship with these measures provided evidence for the concurrent validity of the scale. Fifth, for the predictive validity, a hypotheses testing method was used. Mean scores of PYSÖ showed significant differences for a group of independent variables. According to t-test and variance analyses results, PYSÖ was distinctive for some factors such as age, gender, education level, job, being a mother, experiencing the event in the most affected area, leaving the disaster area immediately, having lost close ones, and being resident in the disaster area at the present time of the study. Sixthly, for the construct validity factor analysis was conducted. It was found that PYSÖ could be used as a two-factor scale as well as a one-factor general scale. Seventh, in terms of internal consistency, satisfactory alpha values showed that the translated Turkish forms of the scale were easily understandable by Turkish groups. Eighth, at the end of the item-total correlation analyses, the correlation values ranged from .78 to .82. Finally, standard error of measurement technique was also used for reliability analyses. Individuals' total scores should be 2 points different. Consequently, it can be concluded that the Turkish form of the Scale has adequate reliability and validity values for use in related research. PYSÖ can be used to evaluate psychological distress in psychometric assessment process.

The present study has several strengths and limitations. Its strengths include the sample size and the use of standardized measures and procedures. The weaknesses were typical of many published studies. The use of retrospective measurement entailed the possibility of inaccuracy of participants' memory of events, due to the time frame between the event and data collection (2 months). Another limitation of the methodology was the use of different measurement methods. Many of the items included in the questionnaire were objective situations or actions. Conversely, some of these scales were largely comprised of subjective ratings of subjective experiences. Replication with different subjects is needed. Replication attempts should involve different stressful events and different populations.

Acknowledgments: This study was supported by a grant from Marmara University, Scientific Research Fund, Istanbul, Turkey. (Grant No. soc-027/020103). The author also wishes to acknowledge Dr. Hobfoll from Kent State University for his invaluable comments during the study.

References

- American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.), Washington, D.C.: Author.
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Klauminzer, G., Charney, D. S., & Keane, T. M. (1990). A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *The Behavioral Therapist*, 13, 187-198.
- Cronbach, L. J. (1951). Coefficient alpha and internal structure of tests. *Psychometrika*, 16, 297-334.
- Derogatis, L. R. (1983). *SCL-90 Administration, scoring and procedure manual-I, for the DSM-R (revised) version*, Baltimore: Johns Hopkins University, School of Medicine.
- Freedy, J. R., & Hobfoll, S. E. (1995). *Traumatic stress: From theory to practice*, New York: Plenum.
- Freedy, J. R., Kilpatrick, D. G., & Resnick, H. S. (1992). Conceptual framework for evaluating disaster impact: implications for clinical intervention. In L. S. Austin (Ed.), *Responding to disaster: A guide for mental health professionals*, (pp. 3-23). Washington D. C.: American Psychiatric.
- Freedy, J. R., Saladin, M. E., Kilpatrick, D. G. Resnick, H. S., & Saunders, B. E. (1994). Understanding acute psychological distress following natural disaster. *Journal of Traumatic Stress*, 7, 257-273.
- Goldberg, D. P. (1972). *The detection of psychiatric illness by questionnaire*, London:Oxford.
- Green, B. L. (1990). Defining trauma: terminology and generic stressor dimensions. *Journal of Applied Social Psychology*, 20, 1632-1642.
- Hobfoll, S. E. (1988). *The ecology of stress*, New York: Hemisphere.
- Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, 44, 513-524.
- Hobfoll, S. E. (1995). Traumatic stress: A theory based on rapid loss of resources. *Anxiety Research*, 4, 187-197.
- Hobfoll, S. E. (1998). *Stress, culture, and community*, New York: Plenum.
- Hobfoll, S. E. (2001). The influence of culture, community, and the nested-self in the stress process: Advancing conservation of resources theory. *Applied Psychology: An International Review*, 50, 337-421.
- Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of event scale: a measure of subjective stress. *Psychological Medicine*, 41, 209-228.
- Karagüven, M. H., & Erktin, E. (2000). *The psychological impacts of Marmara Earthquake in Turkey*, 21st International Conference of the Stress and Anxiety Research Society (STAR), Bratislava.

- Karagüven, M. H., & Erkin, E. (2001). *Olayın etkisi ölçeğinin geçerlik ve güvenilirliği*. [Reliability and validity of impact of event scale] X.Eğitim Bilimleri Kongresinde sunulan bildiri, Abant İzzet Baysal Üniversitesi, Bolu.
- Mirowsky, J., & Ross C. E. (1989). Psychiatric diagnosis as reified measurement. *Journal of Health and Social Behavior*, 30, 11-25.
- Öner, N. & Le Compte, A. (1985). *Durumluk-sürekli kaygı envanteri el kitabı* [Manual for the state-trait anxiety inventory-Turkish form]. İstanbul: Boğaziçi Üniversitesi.
- Öner, N. (2006). Türkiye’de kullanılan psikolojik testlerden örnekler: Bir başvuru kaynağı [Samples of psychology test used in Turkey: A reference material]. Genişletilmiş 2.Baskı, Cilt 2, İstanbul: Boğaziçi Üniversitesi.
- Patterson, G. T., (2005). Development and validation of the psychological distress scale measuring distress, well-being and burden. Retrieved December 19 2008 from <http://sswr.confex.com/sswr/2005/techprogram/P1550.HTM>
- Paulin, C., Lemoine, O., Poirer, L., & Lambert, J. (2000). Validation study of a nonspecific psychological distress scale. *Social Psychiatry-Psychiatry Epidemiology*, 40, 1019-1024.
- Solomon, S. D. (1989). Research issues in assessing disaster’s effects. In R. Gist, and B. Lubin (Eds.), *Psychosocial aspects of disaster*, (pp. 308-340). New York: Wiley.
- Spielberger, C.D., Gorsuch, R. L., & Lusbene, R. E. (1970). *Manual for the state-trait anxiety inventory (self-evaluation questionnaire)*, Palo Alto, CA: Consultant Psychologists.
- World Health Organization. ([WHO], 2008). *World Health Report, WHO Library Cataloguing in Publication Data*, ISBN 978 92 4 0682740. Retrieved December 23 2008 from <http://www.who.int/whosis/whostat/2008/en/index.html>

Psikolojik Yoğun Stres Ölçeği Türkçe Formunun Geçerlik ve Güvenirlik Çalışması

(Özet)

Problem Durumu: İnsanlar yaşamları süresince strese neden olan deprem gibi doğal afetlere, kaza ve terör olaylarına maruz kalabilmektedirler. Bu tür olayların sıklığı ve etkilediği kişi sayısının fazla olması etkilerinin belirlenmesine yönelik çalışmalara önem kazandırmaktadır. Ancak bu alanda yapılacak çalışmalarda kullanılacak geçerli ve güvenilir Türkçe ölçme araçlarının sayısı oldukça sınırlıdır. Araştırmalarda kullanılan ölçme araçlarının yeterli geçerlik ve güvenilirliğe sahip olması ulaşılan bulguların güvenilirliği ile yakından ilgilidir.

Amaç: Bu çalışmanın amacı İngilizce olarak geliştirilmiş olan Psikolojik Yoğun Stres Ölçeği’nin (PYSÖ) Türkçe formunu tanıtmak ve geçerlik ve güvenilirliğine yönelik bulguları değerlendirerek Türkçe’ye kazandırmaktır.

Yöntem: Çalışmada ilişkisel tarama yöntemi kullanılmıştır. Önce ölçeğin Türkçe eşdeğerlik çalışması tamamlanmış daha sonra Marmara

Depreminden sonraki ilk iki ay içinde İstanbul'dan ve deprem bölgesinde çadır kentlerden toplanan veriler üzerinden istatistik analizleri yapılmıştır. Çalışmanın genelinde 1403 kişiye ulaşılmış ancak geçerlik ve güvenilirlik çalışmaları 1144 kişiden toplanan sağlıklı verilerle gerçekleştirilmiştir. Eşdeğerlik sınaması için Boğaziçi ve Marmara Üniversiteleri İngilizce Öğretmenliği son sınıf öğrencilerinden oluşan 100 kişilik bir gruptan yaklaşık 2 hafta arayla uygulamadan elde edilen veriler kullanılmıştır. Daha sonra geçerlik ve güvenilirlik çalışmasına geçilmiştir. 18-55 yaş arası, yetişkinlerden oluşan çalışma grubunun %60'ını erkekler, %40'ını kadınlar teşkil etmektedir. Grubun yaklaşık %63'ü lise, %14'ü orta okul, %9'u üniversite mezunudur. İlave olarak grupta %4 dolayında ilköğretim ve yine aynı oranda lisansüstü eğitimi yapmış ve %0,7 oranında okuma-yazma bilenler yer almıştır. Örneklem; deprem bölgesinden iki büyük çadır kent ve İstanbul'da Anadolu yakası ile sınırlı tutulmuştur. Katılımcılara konuyla ilgili bilgi verilmiş ve katılım gönüllü olmuştur. Çalışmada demografik verileri toplamak için açık ve kapalı uçlu sorulardan oluşan 19 maddelik bir anket formu kullanılmıştır. PYSÖ 9 maddeden oluşan bir belirti tarama listesidir. Yaşanan bir travmadan sonraki 2 ay içinde travmanın etkisini belirlemeye yönelik olarak geliştirilmiştir. Listedeki maddeler bireyin işsel durumunu ve bireyler arası iletişime yönelik yakınmalarını değerlendirmek amacıyla düzenlenmiştir. Bu şikayetler; "Korkunç şeyler düşünmek", "Sinirlilik", "Aşırı üzüntü", "Uyku bozukluğu", "Umutsuzluk", "Kalp çarpıntısı", "Çabuk öfkelenme", "Endişe duyma" ve "Başkalarına karşı sabırsız olmak" ifadeleri ile değerlendirilir. Katılımcı, son günlerde bu şikayetleri hangi yoğunlukta yaşadığını ifade ederek testi doldurur. Şikayetlerini 3 dereceli bir ölçek üzerinde, envanterin başında yer alan yönergeye göre, 0 ile 2 arasında puanlama yaparak bildirir. Puanlamada 0= Hiç, 1=Biraz, 2=Oldukça sık, şeklinde değerlendirilir. Puanlama sonucu 0 ile 18 arasında bir değere ulaşılır. Durumluk Sürekli Kaygı Envanteri (DSKE) toplam 40 sorudan oluşmuştur, 1-4 arasında değer alan 4 dereceli bir ölçektir. İlk 20 madde Durumluk Kaygıyı değerlendirmeye yöneliktir. Her madde Hiç (1), Biraz (2), Çok (3) ve Tamamıyla (4) seçeneklerinden bir tanesinin işaretlenmesi şeklinde uygulanır. İkinci 20 madde ise Sürekli Kaygıyı değerlendirmektedir. Sürekli kaygı ölçeğinin uygulanması sırasında Hemen Hiçbir Zaman (1), Bazen (2), Çok zaman (3) ve Hemen Her Zaman (4) şıklarından bir tanesinin işaretlenmesi gerekir. Olayın Etkisi Ölçeği (OEÖ) 15 madde ve 2 alt boyuttan oluşan bir ölçektir. Alt ölçeklerden birincisi "katılım", ikincisi "kaçınma" boyutuna aittir. 4 dereceli bir ölçektir. Katılımcının okuyarak son günlerde bu ifadeleri ne sıklıkla yaşadığını değerlendirmesi istenir. Uygulama Hiç (0), Nadiren (1), Bazen (3) ve Sıklıkla (5) şıklarından bir tanesinin işaretlenmesi şeklinde gerçekleşir.

Bulgular: İngilizce formun iyi İngilizce bilen gruplara uygulanmasından elde edilen yüksek alfa .79 değeri İngilizce formun Türk öğrenciler tarafından anlaşılır olduğunu göstermiştir. Çeviri işleminden sonra aynı gruba yaklaşık 15 gün sonra Türkçe form uygulanmış ve .64 ve .78 arasında elde edilen yüksek korelasyon değerleri bu iki form arasında benzerlik olduğunu göstermiştir. PYSÖ puanları travma sonrası ortaya

çıkan yoğun stresin göstergesidir. Travma sonrası stresin bir ürünü olarak kaygının ortaya çıktığı bilinmektedir. Kavramsal (uyum) geçerlik testi için geçerliği ve güvenilirliği yüksek ve kullanımı yaygın olan bir kaygı ölçeğinden yararlanılmıştır. DSKE ve PYSÖ' nün aynı gruplara uygulanmasından ulaşılan puanlar arasında .31 ile .41 arasında korelasyonlar belirlenmiştir. Ölçeğin ayırt ediciliğini sınamak amacıyla yapılan hipotez testinde, t-test ve varyans analizi yapılmış ve sonuçta; ölçekten elde edilen puan ortalamalarının yaş, cinsiyet, eğitim düzeyi, meslek ve depremde kaybı olmak gibi demografik değişkenlere bağlı olarak genellikle .001 düzeyinde anlamlı farklılıklar gösterdiği tespit edilmiştir. Yapı geçerliği için açıklayıcı faktör analizi yapılmıştır. İlk analiz sonucunda 6. maddenin (6. Kalp çarpıntısı) 2. faktördeki yükü .10'un altında (9,462E-02) tespit edilmiştir. Bu nedenle 6. madde çıkarılarak faktör analizi ikinci kez tekrar yapılmıştır. Sonuçta toplam varyansın %58.39'unu açıklayan iki faktör tespit edilmiştir. Birinci faktör varyansın %32.7 sini açıklamış ve 5 maddeden oluşmuştur. Bu maddeler "8.Endişe duyma", "1.Korkunç şeyler düşünmek", "5.Umutsuzluk", "3.Aşırı üzüntü" ve "4.Uyku bozukluğu" şeklinde sıralanmış ve ölçeğin "Üzüntü" boyutu olarak adlandırılmıştır. İkinci grup 3 maddeden oluşmuş ve bu maddeler; "7.Çabuk öfkelenme", "2.Sinirlilik ve "9.Başkalarına karşı sabırsız olmak" şeklinde sıralanmıştır. İkinci boyuttaki ifadeler öfke, sinirlilik ve sabırsızlık şeklinde duyguları kapsayan "Heyecan" boyutu olarak adlandırılmıştır. Diğer yönden toplam varyansın (%58.39) yüksek olması ve 1. faktörün döndürmeden önceki faktör yüklerinin yüksek olması (.58-.71) hatta 6.madde çıkarılmadan önce bile faktör yüklerinin yüksek olması (.56-.72) nedeniyle ölçeğin 9 maddelik ve tek boyutlu genel bir ölçek olarak da kullanılabilceği belirlenmiştir. Faktör analizi sonucu bütün maddelerin aynı değişkeni ölçtüğü söylenebilir.

Güvenirlik alfa (.82 ve .83) değerleri ve madde-toplam puan korelasyonlarına (.78 ile .82) bakılarak, iç tutarlık sınaması şeklinde test edilmiştir. Orijinal İngilizce formun alfa değeri Amerika Birleşik Devletlerinde depreme maruz kalan 229 kişilik bir örneklem grubu için .82 düzeyindedir. Bu çalışmada da aynı alfa değeri elde edilmiştir. Güvenirlik sınaması için ayrıca değişik örneklem gruplarında Ölçmenin Standart Hatası hesaplanmış ve 1.98 ile 2.11 arasında değişen değerler elde edilmiştir.

Sonuç ve Tartışma: Bu makalede travmanın psikolojik etkisini değerlendirmeye yönelik olarak kullanılan bir psikolojik yoğun stres ölçeğinin Türkçe formunun psikometrik özellikleri incelenmiştir. Eğitim araştırmalarında geçerlik ve güvenilirlik sınamasına yönelik çalışmalara yaygın olarak rastlanmaktadır. Ancak bu çalışmalar içinde ağır travma sonrası stresi değerlendirmeye yönelik olarak kullanılacak Türkçe ölçeklere ait çalışmaların sayısının oldukça sınırlı olduğu gözlenmektedir. Doğal afet ve kaza gibi stres veren olayların sık yaşandığı ülkemizde, travmaya maruz kalan bireylerin yardıma ihtiyacı olup olmadığını belirlemek amacıyla bu tür ölçeklere ihtiyaç duyulduğu bir gerçektir. Bu

ölçeğin konuyla ilgili alanda kullanımının yaygın olabileceği düşünülmüştür.

Türkiye’de kullanımı çok yaygın olmayan ölçeğin orijinal formu İngilizce olarak geliştirilmiştir. Orijinal form Marmara ve Boğaziçi Üniversitelerinde, yaklaşık 100 kişilik bir gruba uygulanmıştır. Analizler sonucu .79 düzeyinde yeterli alfa değeri elde edilmiştir. Travma sonrası olayın etkisini değerlendirmeye yönelik olarak geliştirilmiş olan 15 maddelik OEÖ ile aynı grup için yazar tarafından yapılan uygulamadan elde edilen alfa değerleri tüm test ve “kaçınma” ve “katılım” alt testlerde .79 ve .89 düzeyinde bulunmuştur. Bu değerler her iki ölçeğinde İngilizce formunun Türk öğrenciler için uygulanabilir olduğunu göstermiştir.

PYSÖ’nin Türkçe ve İngilizce Formları Puanları Arasındaki Korelasyon ölçeğin Türkçe ve İngilizce formları arasında paralellik olduğunu gösterecek şekilde .77 dir. Ölçeğinin Türkçe formunun uygulanmasından elde edilen puan bir olay sonrası ortaya çıkan stresin göstergesidir. Elde edilen ölçümlerin doğruluğu kavramsal (uyum) geçerliği ve yapı geçerliği yöntemleriyle sınanmıştır. Stres ve kaygı bir arada bulunduğundan, geçerlik sınaması için, kaygıyla ilgili çalışmalarda yaygın olarak kullanılan ve tatminkar psikometrik değerlere sahip olan DSKE kullanılmıştır. DSKE ile PYSÖ puanları arasında oldukça yüksek korelasyonlar bulunmuştur. PSYÖ’nin ayırt ediciliğini test etmek için t-test analizi yapılmış ve sonuçta; depremden çok etkilenen bölgede (çadır kentlerde) yaşayanlar ile daha az hissedilen bölgede (İstanbul, Anadolu yakasında) yaşayanlar, olaydan sonra bölgeden hemen ayrılanlar ile ayrılmayanlar, olayda bir yakınına kaybedenler ile kaybetmeyenler ve çocuk sahibi olan kadınlar ile olmayanlar arasında ayırt edici olduğu belirlenmiştir. Varyans analizi sonucunda ise yaş, eğitim düzeyi ve halen deprem bölgesinde yaşıyor olmak gibi faktörlerin puan ortalamaları üzerinde anlamlı farklılığa neden olduğu ulaşılmış önemli bulgulardandır. Ölçeğin ayırt ediciliği yüksektir. Yapı geçerliğini sınamak için yapılan faktör analizi sonucunda ölçeğin “üzüntü” ve “heyecan” olarak adlandırabilecek iki boyuttan oluştuğu belirlenmiştir. Ancak aynı zamanda bütün maddelerin aynı değişkeni ölçtüğü ve maddelere verilen cevapların genel bir ölçek puanı olara toplanabileceği görülmüştür. PYSÖ puanı yaklaşık olarak 2 puan hata farkı gösterebilmektedir. Sonuç olarak ölçek Türkçe’ye kazandırılmış ve yeterli psikometrik değerlere ulaşılmıştır.

Öneriler: PYSÖ çalışmalarda psikolojik yoğun stresi değerlendirmek amacıyla kullanılabilir. Benzer çalışmaların tekrarı gereklidir. Tekrarlar değişik stres verici durumlarda ve farklı çalışma grupları ile gerçekleştirilmelidir.

Anahtar Sözcükler: Psikolojik yoğun stres ölçeği, geçerlik, güvenilirlik

The Effects of an Emotional Intelligence Skills Training Program on the Emotional Intelligence Levels of Turkish University Students

T.Fikret Karahan**

B. Murat Yalçın**

Suggested Citation:

Karahan T.F., Yalçın B.M. (2009). The effects of an emotional intelligence skills training program on the emotional intelligence levels of Turkish university students. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 36, 193-208.

Abstract

Problem Statement: Emotional intelligence is the capacity to require as well as apply information of an emotional nature and to feel and to respond emotionally. Emotional intelligence contains five main domains as follows: self-awareness, managing emotions, motivating oneself, empathy, and handling relations. Individuals who are emotionally competent are at an advantage in every area of life, whether family and relations, education, community, and organizational pursuits, and are more likely to lead happy and productive lives. As emotional intelligence may be developed and learned at any time or any age, one of the main gaps in this area is the lack of effective programs for Turkish adults. In order to promote the emotional intelligence of Turkish adults in the short and long term, we designed an "Emotional Intelligence Skills Training Program."

Purpose of the Study: To determine whether individuals would improve their emotional intelligence and, if so, whether the beneficial effects of the program are maintained over time.

Methods: Following the announcement that an "Emotional Intelligence Program," which is intended to improve emotional intelligence skills, was to be held at the Education Faculty of Ondokuz Mayıs University, the Emotional Intelligence Self-Evaluation Scale was administered to 215 university students who volunteered to participate from different faculties. The 20 group members with the lowest test scores from each gender were randomly selected (40 group members in total) and randomized into a study group and a control group (10 males and 10 females each). A program of 12 main sessions and two extra sessions of 90 minutes (three and six months after the end of the program) were applied to the members of the study group. The scales were re-administered to the study group and the control group at the end of the program (post-test) and to the study group just after the extra sessions (first and second follow-up tests).

* Asst. Prof. Dr., Ondokuz Mayıs University Faculty of Education, tfikretkarahan@hotmail.com

** Asst. Prof. Dr., Ondokuz Mayıs University Faculty of Medicine, myalcin@omu.edu.tr

Findings and Results: Although the scores of the two groups were no different before the commencement of the program ($P > 0.5$), the study group had better result scores at the termination of the program compared to those of the control group ($P < 0.001$). The study group's scores remained improved three and six months after the end of the program ($P = 1$).

Conclusions and Recommendations: We determined that the university students admitted into this program improved their emotional intelligence skills in terms of the inventory scores over time.

Keywords: Emotional intelligence, emotions, adult, intervention, university students, psychotherapy, Turkey

Social scientists have been trying to categorize and define intelligence for many years. Until a few decades ago, intelligence was measured only in terms of IQ, which attempts to tap the individual's cognitive capacity and functioning (e.g., the ability to learn, recall, apply, think, reason, and abstract). Early insights resulted from David Wechsler's (Kaufman, A. S. & Kaufman, 2001) idea of "non-intellectual aspects of general intelligence" followed by Leeper's (1948) proposal that "emotional thought" is a part of and a contribution to "logical thought" and intelligence in general. Later, Gardner (1983) indicated that intelligence encompasses multiple dimensions, combining a variety of aspects with "personal intelligence," which he referred to as "intrapersonal capacities" and "interpersonal skills" as the two main components. After Mayer and Salovey (1997) developed their model based upon Gardner's ideas, the concept gained popularity with the work of Daniel Goleman (1995; 1998; 2000).

Finally, Bar-On (1997; 2000; 2007) defined emotional intelligence as the sum of incognitive abilities concerned with understanding oneself and others, relating to other people and adapting to and coping with one's immediate surroundings in order to be more successful in dealing with environmental demands. He defines emotional intelligence as tactical (immediate functioning) and cognitive intelligence as strategic (long-term capacity). Under the Bar-On model, emotional intelligence has 15 components derived from the five basic domains. The first of these is the intrapersonal domain, which denotes self-awareness and control. It is made up of self-regard, emotional self-awareness, assertiveness, independence, and self-actualization. The interpersonal domain denotes the ability to form relationships and is made up of empathy, social responsibility, and interpersonal relationships. The adaptability domain represents the ability to be flexible yet realistic and the ability to solve problems, consisting of reality testing, flexibility, and problem solving. Stress management components include the abilities needed in order to overcome stress denotes, such as the ability to route stress in order to minimize its effects (stress tolerance, impulse control, etc.). Finally, the general mood component denotes the ability to accept one's self and others as they are and to enjoy life as it is. It is made up of optimism and happiness.

As individuals with effective emotional intelligence skills form more healthy emotional and social relationships and develop a positive quality of life, it is not surprising that efforts have concentrated on improving these skills (Eack et al., 2007; Wiegand, 2007; Akerjordet & Severinsson 2007; Codier, Kooker, & Shoultz, 2008).

However, insufficient data exists on this subject in Turkey. One of the main gaps in this area is the lack of effective programs for Turkish adults. In order to promote the emotional intelligence of Turkish adults in the short and long term, we designed an "Emotional Intelligence Skills Training Program," employing a Turkish cultural perspective and accumulated data. The aim of this study was to determine whether individuals would improve their emotional intelligence and, if so, whether the beneficial effects of the program are maintained over time.

Method

Research Design

In order to investigate the effectivity of The Emotional Intelligence Skills Training Program, a pretest, posttest, and follow-up study with a control group consisting of volunteered university students was designed.

Sample

The control and study group consisted of 40 volunteer university students of Ondokuz Mayıs University, Samsun.

Procedure

After it was announced that an "Emotional Intelligence Skills Training Program" aimed at improving emotional intelligence was to be held at the Ondokuz Mayıs University Education Faculty between 1 March and 1 May, 2006 in Samsun, Turkey, a total of 215 university students from different faculties (23 Agriculture, 16 Dentistry, 47 Arts of Science and Education, 115 Education, and 14 Engineering) voluntarily applied to participate in the program. The Emotional Intelligence Self-Evaluation Scale (Hall, 1998) was administered to these students. The results of the male and female group members were pooled separately. The 20 group members with the lowest test scores from each gender were randomly selected (40 group members in total) and divided into a study group and a control group, consisting of 20 (10 males and 10 females) individuals. The study group then attended the "Emotional Intelligence Skills Training Program" over 12 weeks with the meetings lasting 90 minutes every weekend, whereas the control group was put on a waiting list (the individuals in the control group participated in the Emotional Intelligence Skills Training program just after the study group's program ended). The other group members were also put on the waiting list. Each individual was asked to attend all the sessions punctually, and all agreed to do so. At the end of the program, the scale was again applied to the study and control groups. Three and six months after the termination of the program, two extra sessions were held in order to review the effects of the program and discuss how it was reflected in group members' lives. Just after the extra session (90 minutes), the scales were re-administered. Afterwards, the pre-test, post-test, and follow-up test scores of the study and control groups were compared. None of the group members dropped out of the study.

Research Instruments

Emotional Intelligence Self-Evaluation Scale. Nicholas Hall (1998) developed the Emotional Intelligence Self-Evaluation Scale. The scale is suitable for individuals over 18 years of age, contains 30 items in the form of short sentences, and employs a 6-point response scale with a textual response format ranging from "Disagree very much" [1] to "Agree very much" [5].

The scale is a self-reporting measure of emotionally and socially intelligent behavior that provides an estimate of emotional-social intelligence. It is designed to measure a number of constructs related to emotional intelligence and gives an overall emotional intelligence score as well as scores for the following five composite scales: (1) emotional awareness, (2) managing one's emotions, (3) self-motivation, (4) empathy, and (5) coaching others' emotions. Scores for the scale will almost always be between 30 and 180. Average (between 130-154 points) to above average (>155) emotional intelligence scores on the scale suggest that the respondent is effective in emotional and social functioning. The higher the scores, the greater the likelihood of effective functioning in meeting daily demands and challenges. On the other hand, low emotional intelligence scores (≤ 129 points) suggest an inability to be effective and the possible existence of emotional, social, and/or behavioral problems. The inventory was adapted and translated into Turkish by Yılmaz and Ergin (2000). While adapting the scale in Turkish, the original form of the scale was administered to 55 university students at the English Department of Foreign Languages and Literature Faculty who were in their final year. All had achieved a minimum mark of 75 (out of 100) in the previous year. The scale was then translated into Turkish with a group of academics from the English Department of the Foreign Languages and Literature Faculty and Department of Psychological Counseling and Guidance from the Education Faculty. This group discussed every item of this scale and translated them all into Turkish. The translated scale was re-administered to the same student sample a month from the application of the first test. Items, subscales, and total scores of the English and Turkish versions of both scales were compared and correlated. The Cronbach-alpha values for the confidence interval for the first 30 items between the two applications lay between 0.76 and 0.92. The first subscale correlation between two applications was 0.76, followed by the second subscale at 0.92, third subscale at 0.91, fourth subscale at 0.86, and, finally, fifth subscale at 0.84. These results statistically indicated that the English and Turkish versions of the scales were similarly very well perceived and understood by the group members. The item analyses from the total scores in the Turkish version and the relation between each item and the total test score were investigated. The relation coefficients calculated in this way (except for two items) varied between 0.73 between 0.89. The items in the score were separately investigated as odd and even numbers, and the correlation coefficient was 0.79.

Emotional Intelligence Skills Training Program. The independent variable of this research was the "Emotional Intelligence Skills Training Program." In designing the

program, the authors followed an eclectic approach, referred to a large number of sources, and took cultural factors into consideration (Goleman, 2005; Davis, 2004; Erkan, 2004; Blankstein, 2003; Altınay, 2003; Cooper & Sawaf, 2003; Konrad & Hendl, 2003; Kulaksızoglu, 2003; Selçuk & Güner, 2003; Stein & Book, 2000; Güney, 2000; Çetin et. al., 2001; Brockert & Braun, 2000; Goleman, 2000; Damasio, 1999; Goleman, 1998; Shapiro, 1998; Weisinger, 1998; Dökmen, 1997; Baltaş & Baltaş, 1992). The program has an educative and time-limited structure based on the small-group experience. The main aim of the program is to improve group members' emotional intelligence. The program covered the areas of being aware, identifying, perception, differentiating between emotions, being aware of the methods of expressing emotions, understanding the relationship between emotions and thoughts as well as physical reactions and behavior, managing emotions, displaying empathic bonding with others and empathic reactions to achieve empathic listening skills, learning to expend motivational energies in the direction of a determined target and in a specific way, using motivating and trusting speech, differentiating between behavior that is friendly and that which is not, forming positive thoughts in friendships, being aware of the presence of multiple solutions to a specific problem, and developing skills for the management of relationships. The program also includes relaxation training.

The sessions were administered with the assistance of the researchers (Karahan & Yalcin). Sessions included such measures as the provision of skill-related information, role-playing, and scenarios based on real or fiction-based experience and homework. An assessment was carried out together with the group members after each session, and they were assigned exercises to be performed at home with the aim of encouraging them to apply the information obtained in class to their day-to-day lives. The "Emotional Intelligence Skills Training Program" schedule is presented in Appendix 1.

Statistical Analyses. The numerical results of the Emotional Intelligence Self-Evaluation Scale were regarded as dependent variables in this study. The Dependent t-test was used to investigate the relation between the pre-test and post-test scores of the study and control groups and for comparisons within the groups. The differences between the pre-test, post-test, and follow-up test results of the study group were examined using The Variance Analyses in Dependent Groups (Repeated Measures), and Bonferroni's test (the *t-test* in Dependent Groups) was used to investigate the relationship between the pair results. The effect of the group member's gender on these tests was also investigated using within a subjects factor. $P < 0.05$ was regarded as significant. All analyses were performed using SSPS 13.0 (SSPS Inc., Chicago, IL).

Findings and Results

There was no statistically significant difference between the two groups in terms of age (study group=22.06±4.43, control group=22.22±5.2 years) and educational status (study group=13.32±0.9, control group=13.24±0.5 mean years spent in

education) ($P > 0.5$). The scores of the study group were pre-test=102.95±5.81, post-test=142.75±9.71, first follow-up=142.75±9.86, and second follow-up=142.60±9.37, and those of the control group were pre-test=102.80±5.29 and post-test=102.65±5.22. In light of the pre-test results, the Emotional Intelligence Self-Evaluation Scale revealed that the group members had low to moderate levels of emotional intelligence skills ($P > 0.5$).

The initial results of the Emotional Intelligence Self-Evaluation Scale applied to the control and study groups revealed no significant difference ($t=0.085$, $P=0.933$). However, post-test scores (at the end of the program) for those attending the program were significantly better than their initial scores ($t=17,888$, $P<0.001$) and than the scores of those group members who did not attend ($t=16,263$, $P<0.001$). The study group's post-test, three-month follow-up, and six-month follow-up Emotional Intelligence Self-Evaluation Scale scores were significantly better than the pre-test scores ($F= 316,569$, $P< 0.001$). There were no significant differences between post-test and first follow-up scores ($P= 0.1$), post-test and second follow-up scores ($P=0.1$), and first follow-up and second follow-up ($P= 0.1$) in pair comparison (Bonferroni test).

Conclusions and Recommendations

We determined that the university students admitted into this program improved their emotional intelligence skills in terms of the inventory scores. The fact that the study group's scores remained better than their initial scores after six months may be more significant. According to Weisinger (1988) and Stein and Book (2000), being aware, understanding, and expressing a feeling fully and correctly seems to be the key element in self-understanding and change together with the management of emotions. We, therefore, intended that our subjects should become more aware and better able to perceive, to define, to discriminate, to think, to notice, and to understand their emotions. We also intended that our group members should improve the expression of their emotions, the management of them, the comprehension of the relation between emotions, thoughts, and reactions, and the establishment of empathic communications. Motivation was an important domain of the program as the group members were coached to gain motivational energies and direct these toward a determined objective, to communicate in a motivating manner, to form positive thoughts about their experiences, to understand verbal and non-verbal behavior that may facilitate consensus in relationships, to differentiate between friendly and hostile behavior, to manage relationships, and to empathize with others. Improving the domains of emotional intelligence (self-awareness, self-coping skills, self-motivation, management of emotions, effective communication skills, and empathy) undoubtedly helps individuals better manage emotional and social problems, which may arise from their daily lives, form healthier relationships, and, enjoy life.

Although the program ran successfully and smoothly, there were some unforeseen problems over which the researchers had to intervene. One problem was to obtain contribution to the group process from all the group members. This problem was not fully resolved till the second session. Some group members had some difficulties in abiding by the group rules, which were determined at the beginning of the first session. Some were very active and had a tendency to interrupt communication. They preferred to communicate with the therapist directly, ignoring the group process in the first two sessions, although the group process neutralized and reduced this behavior. At the beginning of the program, some group members found the role-playing and scenarios childish and hesitated to assume a role in these; they had to be motivated with group support. In the role-playing, some of the group members had difficulties expressing basic or complicated emotions. They stated that although they understood those feelings, they did not know how to express them (offense, disappointment, restlessness, imitation, yearning, apprehension, etc.). Generally, all the group members had difficulties defining their emotions. The main difficulties came in identifying and differentiating their emotions from their thoughts and ideas, which is obvious in complicated emotions. In addition, the group members were well able to identify and be aware of negative emotions, while they had problems with positive ones. Although the group members were aware of the importance of empathy, the empathic skills of both male and female group members were determined to be highly inadequate; they seemed to prefer acting in an egocentric manner. The researchers employed several empathy-building techniques in a variety of scenarios in order to improve the group members' skills.

However, this study may have some limitations. First, although the aim of this study was to improve the group members' emotional intelligence, the outcome of the program was investigated in the light of the scores of the Emotional Intelligence Self-Evaluation Scale, which does not directly reflect the positive changes in the group members' lives. All such feedback was obtained through subjective methods. The selection of the group members may have had an effect on our results as they had the lowest emotional intelligence skills of all the participants. In addition, they may have been more highly motivated as normal individuals, since no one dropped out of the program or even missed a single session. It may not, therefore, be possible to generalize our results to individuals with moderate or high levels of emotional intelligence. Furthermore, although highly experienced therapists ran the program (Karahan & Yalcin, 2009), the success of similar programs was highly dependent on the therapists' abilities and skills as no observer assessed their performance objectively.

In conclusion, further research is needed with group members from different socioeconomic backgrounds, demographic features, and levels of emotional intelligence in order to investigate the effect of this program on their skills, and different methods need to be employed in order to investigate the reflection of these skills in their daily lives.

References

- Akerjordet K., & Severinsson E. (2007). Emotinal intelligence: A review of the literature with specific focus on empirical and epistemological perspectives. *Journal of Clinical Nurse*, 16(8), 1405-1416.
- Altınay, D. (2003). *Psikodrama. 400 ısınma oyunu* [psychodrama: 400 warming games]. Istanbul: Sistem Press, 4th. Edition.
- Bar-On, R. (1997). *The Emotional Quotient Inventory (EQ-I): A test of emotional intelligence*. Toronto, Canada: Multi-Health Systems Inc.
- Bar-On, R. (2000). *Emotional and social intelligence: Insights from the Emotional Quotient Inventory (EQ-I)*. In R. Bar-On and J. D. A. Parker (Eds.), *Handbook of emotional intelligence*. San Francisco: Jossey-Bass.
- Bar-On, R. (2004). The Bar-On Emotional Quotient Inventory (EQ-i): Rationale, description, and summary of psychometric properties. In Glenn Geher (Ed.), *Measuring emotional intelligence: Common ground and controversy*. Hauppauge, NY: Nova Science Publishers.
- Bar-On, R. (2006). *The Bar-On model of emotional-social intelligence (ESI)*. *Psicothema*, 18, suppl.
- Bar-On, R. (2007). *The Emotional Quotient Inventory (EQ-I): Technical manual*. Toronto, Canada: Multi-Health Systems Inc.
- Baltaş, Z. & Baltas, A. (1992). *Bedenin dili* [the body language]. İstanbul: Remzi Press, 3rd Edition.
- Baltaş, A. (2001). *Değişimin içinden geleceğe doğru ekip çalışması ve liderlik* [inside change teamwork and leadership towards future]. Istanbul: Remzi Press, 2nd Edition.
- Blankstein, A. M. (2003). *EQ+IQ= Best leadership practices for caring and successful schools*. In Maurice J. Elias, Harriett Arnold, and Cynthia Steiger Hussey (Eds.), New York, USA: Corwin Press Inc.
- Brackett, M. A., Mayer J. D. & Warner, R. M. (2003). Emotional intelligence and its realiton to everyday behaviour. *Personality And Individual Differences*, 36, 1382-1402.
- Brockert, S. & Braun, G. (2000). *Using your emotional intelligence*. Nurettin Süleymangil (Trans.). Istanbul: MNS Press.
- Codier, E., Kooker, B. M. & Shoultz, J. (2008). Measuring the emotional intelligence of clinical staff nurses. An approach for improving the clinical care environment. *Nurse Admission Quality*, 32(1), 8-14.
- Cooper, R. K. & Sawaf, A. (2003). *Liderlikte duygusal zeka* [emotional intelligence in leadership]. Ali Varlı (Trans.). Istanbul: Sistem Press.
- Cetin, F. & et.al. (2001). *Araştırmadan uygulamaya çocuklarda sosyal beceriler* [from research to application: Social skills in children]. Istanbul: Epsilon Press.

- Damasio, A. R. (1999). *Descartes'in yanılması: Duygu, akıl ve insan beyni*. [Descartes' error: Emotion, cognition and the human brain]. Bahar Atlamaz (Trans.), İstanbul: Varlık/Bilim Press, 2nd Edition.
- Davis, M. (2004). *EQ, duygusal zekanızı ölçün* [EQ, measure your emotional intelligence]. Solina Silahlı (Trans.). İstanbul: Alfa Press.
- Dökmen, Ü. (1988). Empatinin yeni bir modele dayanılarak ölçülmesi ve psikodrama ile geliştirilmesi [measuring empathy according to a new model and development via psychodrama]. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 21(1), 175-189.
- Dökmen, Ü. (1997). *İletişim çatışmaları ve empati* [conflicts of communication and empathy]. İstanbul: Sistem Press, 5th Edition.
- Eack, S. M., Hogarty, G. E., Greenwald, D. P. & et al. (2007). Cognitive enhancement therapy improves emotional intelligence in early course schizophrenia: Preliminary effects. *Schizophrenia Research*, 89(1-3), 308-311.
- Erkan, S. (2004). *Örnek grup rehberliği etkinlikleri* [examples of group guidance activities]. İstanbul: Pegem A Press, 6th Edition.
- Goleman, D. (1998). *Duygusal zeka neden IQ'dan daha önemlidir?* [why is emotional intelligence more important than IQ?]. Banu S.Yüksel (Trans.), İstanbul: Varlık/Bilim Press, 7th Edition.
- Goleman, D. (2000). *İşbaşında duygusal zeka* [emotional intelligence on work]. İstanbul: Varlık Press. 3rd Edition.
- Goleman, D. (2005). *Yıkıcı duygularla nasıl başa çıkabiliriz?* [how can we overcome destructive emotions]. Seda Hauser (Trans.). İstanbul: İnkılap Press.
- Güney, S. (2000). *Davranış bilimleri* [behavioral sciences]. 2nd Edition. Ankara: Nobel Press.
- Hall, N. (1998). Can emotional intelligence be learned? *Harvard Business Review*, 176(6), 93-105.
- Karahan, T. F. & Yalcin, B. M. (2009) The effects of an emotional skills training program on the anxiety, burn-out and HgA1c levels of type 2 Diabetes Mellitus patients. *Türkiye Klinikleri Tıp Bilim Dergisi*, 29(1), 16-24.
- Kaufman, A. S. & Kaufman, J. C. (2001). Emotional intelligence as an aspect of general intelligence: What would David Wechsler say? *Emotion*, 1(3), 258-264.
- Kulaksızoğlu, A. (2003). *Kişisel gelişim uygulamaları*. [applications for personal development]. İstanbul: Nobel Press.
- Kocayörük, A. (2004). *Duygusal zeka eğitiminde drama etkinlikleri*. (drama activities for education in emotional intelligence). Ankara: Nobel Press.
- Leper, R.W. (1948). A motivational theory of emotions replace "emotions disorganized responses". *Psychological Review*, 55, 5-21.
- Konrad, S. & Hendl, C. (2003). *Duygusal zeka sayesinde başarılı bir hayat. Duygularla güçlenmek* [a successful life with emotional intelligence: To be empowered by emotions]. Meral Taştan (Trans.). İstanbul: Hayat Press.

- Matthews, G. & Zeidner, M. (2000). *The handbook of emotional intelligence: Theory, development, assessment and application at home, school and in the workplace*. Reuven Bar-On, James D. A. Parker (Eds.). San Fransisco, USA: Jossey-Bass A Wiley Company.
- Mayer, J. D., Salovey, P. (1997). *What is emotional intelligence? Emotional development and emotional intelligence*. New York, USA: Basic Boks.
- Mayer, J. D., Salovey, P. & Caruso, D. (2000). *Models of emotional intelligence, handbook of intelligence*. Cambridge, USA: Cambridge Universty Press.
- Mayer, J. D., Salovey, P., Caruso, D. R. & Sitarenios, G. (2001). Emotional intelligence as standard intelligence. *Emotion*, 1(3), 232-242.
- Mayer, J. D., Salovey P. & Caruso, D. R. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, 15(3), 197-211.
- Pibermik-Okanovic, M., Peros, K., Szabo, S., Begic, D. & Metelko, Z. (2005). Depression in Croatian type 2 diabetic patients: Prevalence and risk factors. A croatian survey from the European depression in diabetes research consortium. *Diabetic Medicine*, 22, 942-945.
- Savasır, I. & Sahin, N. H. (1997). *Bilişsel-davranışçı terapilerde değerlendirme: Sık kullanılan ölçekler* [evaluation in cognitive behavioral therapies: Commonly used scales]. Ankara: Türk Psikologlar Derneği Yayınları. No: 9.
- Selcuk, Z. & Güner, N. (2003). *Sınıf içi rehberlik uygulamaları*. [guidance applications in class]. İstanbul: Pegem Press., 3rd Edition.
- Shapiro, L. E. (1998). *Yüksek EQ'lu bir çocuk yetiştirmek*. [to raise a child with a superior EQ]. Ümran Kartal (Trans.). İstanbul: Varlık Press.
- Snoek, F. J. & Skinner, C. (2002). Psychological counselling in problematic diabetes: Does it help?. *Diabetic Medicine*, 19(4), 265-273.
- Stein, S. J & Book, H. E. (2000). *EQ duygusal zeka ve başarının sırrı* [EQ emotional intelligence and the secret of success]. Müjde Işık (Trans.), İstanbul: Özgür Press.
- Weisinger, H. (1998). *İş yaşamında duygusal zeka* [emotional intelligence in business life]. Nurettin Süleymangil (Trans.). İstanbul: MNS Press.
- Wiegand, D. M. (2007). Exploring the role of emotional intelligence in behaviour -based safety coaching. *Journal of Safety Research*. 38(4), 391-398.
- Yılmaz, M. (2002). *Duygusal zeka düşünme becerileri eğitiminin annelerin duygusal zeka düzeyine etkisi*. [The effect of an emotional intelligence skills training on the emotional intelligence levels of mothers]. Unpublished doctoral dissertation, Ondokuz Mayıs Üniversitesi, Sosyal Bilimler Enstitüsü, Samsun.
- Yılmaz, H. & Ergin, E. (2000). *Üniversite öğrencilerinin sahip oldukları duygusal zeka düzeyleri ile 16 kişilik özelliği arasındaki ilişki üzerine bir araştırma*. [A study of the relation between the emotional intelligence levels of the university students with thier 16 personality traits]. VIII. Ulusal Eğitim Bilimleri Kongresi Sunulmuş Bildiri, Trabzon.

Bir Duygusal Zeka Becerileri Geliştirme Programının Üniversite Öğrencilerinin Duygusal Zekaları Üzerindeki Etkisi

(Özet)

Problem Durumu: Sosyal bilimciler yıllarca zeka tanımı yaparken zekanın bilişsel ve rasyonel özellikleri üzerine odaklanmışlardır. Bu bağlamda bireyin öğrenme, hatırlama, uygulama, mantık, nedenleme, hesaplama yetileri zekanın ana bileşenleri olarak görülmüştür. Oysa son zamanlarda duyguların davranışlara yön verdiği, değerleri şekillendirdiği ve başkalarına karşı olan davranış biçimlerini seçmede bireyi etkilediği gösterilmiştir. Günümüzde duygusal (EQ) ve rasyonel (IQ) bilişsel becerilerin ikisi de insan zekasının birbirine muhtaç, eşit önem arz eden parçaları olarak kabul görmektedir. Duygusal zeka; duygusal doğadaki bir bilgiyi alabilme ve değerlendirme, duygusal karşılık verme ve hareket etme kapasitesi olarak tanımlanmaktadır.

Duygusal zekayı irdelemek için çeşitli modeller önerilmiş olsa da çoğu kez kuramcılar duygusal zekayı oluşturan, kişisel ve kişilerarası ana bölümleri şekillendiren beş alt öge konusunda hem fikir gözükmemektedirler. Bunlar; öz farkındalık, empati, duygusal yeterlilikler, motivasyon ve ilişki yönetimi olarak sıralanabilir. Bu öğelerden öz farkındalık; bireyin benzerlikleri, farklılıkları, umutları, tercihleri, kültürel mirası, doğal yetenekleri, eksikleri ve bireyi birey yapan diğer eşsizlikleri bilmesi, içsel ve dışsal durumlardan ve süreçlerden haberdar olması şeklinde tanımlanabilir. Empati; diğer kişilerin bakış açılarını ve duygularını anlama, bunlara yönelik ilgili ve şefkatli tutumları geliştirme olarak tanımlanabilir. Duygusal yeterlilikler; duygular için bir birikim oluşturma, duygu-düşünce-davranış arasındaki ilişkiyi kavrama, başka bireylerin duygusal ipuçlarını anlama ve ayırt etme, onlara yanıt verme, stres kontrolü, derinde yatan duyguları anlama, duyguları yapıcı bir şekilde kontrol ve ifade etmeyi öğrenme olarak tanımlanabilir. Motivasyon; bir görev ya da hedef için içsel odaklı bir arzunun olumlu yönlendirici gücü olarak tanımlanabilir. İlişki yönetimi ise aktif dinleme, uygun iletişim teknikleri (ben dili kullanımı, uygun geri bildirim verebilme v.b), problem çözme, çatışma çözme gibi becerileri kapsamaktadır.

Duygusal olarak yetenekli, kendi duygularını denetleyen, başkalarının duygularını tanıyıp etkin bir şekilde yanıt veren bireyler duygularını daha başarılı biçimde yönetebilmekte, duygusal sorunların çözümünde ve stres yönetiminde daha başarılı olmakta, bununla bağlantılı olarak aile içi ilişkilerde ve sosyal ilişkilerde daha yapıcı ve pozitif tepkiler sergilemektedirler. Diğer yandan duygusal zeka düzeyi yüksek olan bireyler, problemlerin çözümünde etkili başa çıkma becerilerine daha fazla

sahip olup, duygusal farkındalık ve duyguların kontrolü konusunda daha başarılıdırlar. Duygusal zeka düzeyi düşük olan bireyler ise, sosyal ilişkilerde daha başarısız olmakta ve daha fazla saldırgan davranışlar sergileyerek olumsuz ilişkiler geliştirmektedirler. Duygusal zekanın öğrenilmiş alışkanlıklar temelinde dayandığı görüşü, uzmanları bu alandaki kapasitenin geliştirilmesinde duygusal zeka eğitiminin rolünü ve önemini incelemeye yöneltmiştir. Araştırmacılar, duygusal zeka eğitiminin, bireylerin duygusal ve sosyal alanlarda daha sağlıklı ilişki kurmalarını, yaşam kalitelerini olumlu yönde geliştirmelerini amaçladığını rapor etmektedirler. Yapılan çalışmalar duygusal zekanın her zaman ve yaşta öğrenilebilir olduğunu göstermiştir. Ülkemizdeki bu konudaki en büyük eksikliklerden birisi yetişkinler için uyarlanmış, bu bireylerin duygusal zeka becerilerini geliştirmeye yönelik programların olmamasıdır. Bu araştırmada araştırmacılar tarafından literatürden yararlanılarak geliştirilen oniki oturumluk Duygusal Zeka Beceri Eğitimi Programı'nın, üniversite öğrencilerinin duygusal zeka düzeyleri üzerindeki etkisi incelenmiştir.

*Araştırmanın Amacı:*Araştırmada deney grubuna verilen duygusal zeka beceri eğitiminin etkilerinin kısa ve uzun süreli olarak kalıcı olup olmadığının, kontrol grubu ile karşılaştırmalı olarak araştırılması amaçlanmıştır.

*Araştırmanın Yöntemi:*Ondokuz Mayıs Üniversitesi Eğitim Fakültesinde katılımcıların duygusal zeka becerilerini arttırmaya yönelik "Duygusal Zeka Becerileri Eğitim Programı" adlı bir kurs düzenleneceği ilan edildikten sonra bu çalışmaya katılmak için çeşitli fakültelerde öğrenim gören 215 öğrenci başvurmuştur. Başvuran öğrencilere kursun amaçları ve uygulanışı hakkında kısaca bilgi verilmiştir. Daha sonra tüm gönüllülere Duygusal Zeka Öz Değerlendirme Ölçeği uygulanmıştır. Her bir cinsiyetten ölçekten en düşük skora sahip 20 grup üyesi seçilmiş (toplam 40 grup üyesi) ve bu yine eşit cinsiyette çalışma ve kontrol gruplarını rastgele yerleştirilmiştir. Doksan dakikalık 12 ana oturumdan oluşan Duygusal Zeka Becerileri Geliştirme Programı çalışma grubuna uygulanmıştır. Program bittikten sonra üçer ay ara ile iki adet toparlayıcı ve hatırlatıcı oturum çalışma grubunda tekrarlanmıştır. Çalışma ve kontrol gruplarının programa başlamadan hemen önceki ve program sona erdikten hemen sonraki skorları kendi içlerinde ve diğer grupla kıyaslanmıştır. Çalışma grubunda toparlayıcı ve hatırlatıcı her bir oturumdan hemen sonra (program bittikten üç ve altı ay sonra) birer adet ölçek izleme amaçlı uygulanmış (Birinci ve ikinci izleme testleri) ve çalışma grubunda bu skorlar grup içi değerlendirilmiştir. Verilerin analizinde; Dependent t-test, The Variance Analyses in Dependent Groups (Repeated Measures) ve Bonferroni's test (The t-test in Dependent Groups) teknikleri kullanılmış olup, SSPS 13.0 paket programından yararlanılmıştır. Sonuçların yorumlanmasında ise $P < 0.05$ anlamlılık düzeyi üst değer olarak alınmıştır.

Araştırmanın Bulguları: Program öncesinde deney ve kontrol gruplarının Duygusal Zeka Değerlendirme Ölçeği ön-test puanları arasında anlamlı bir farklılığın bulunmadığı saptanmıştır ($P>0.5$). Program sonrasında; deney grubunun ön-test ve son-test puanları arasındaki farkın $P< 0.001$ düzeyinde anlamlı olduğu saptanmıştır. Ayrıca deney ve kontrol gruplarının son-test puanları arasındaki farkın $P< 0.001$ düzeyinde anlamlı olduğu saptanmıştır. Deney grubunun son-test puanları ile birinci ve ikinci izleme testi puanları arasında ise anlamlı bir farklılığın bulunmadığı saptanmıştır ($P= 1$).

Araştırmanın Sonuçları ve Öneriler: Bu çalışma sonucunda programa alınan üniversite öğrencilerinin duygusal zeka düzeylerinin program sonrasında arttığı ve üçüncü ve altıncı aylardaki ölçümlerde bu olumlu etkinin devam ettiği izlenmiştir. Bu programın etkinliğinin, değişik yaş gruplarında ve farklı sosyo-ekonomik düzeydeki deneklerle farklı çalışmalarla araştırılıp desteklenmesi gerekmektedir.

Anahtar Sözcükler: Duygusal zeka, duygular, erişkin, uygulama, üniversite öğrencisi, psikoterapi, Türkiye

Appendix 1. The Emotional Intelligence Skills Training Program Schedule

1st Session

Aim: The first session consisted of meeting the trainees, providing information on the Emotional Intelligence Skills Training Program, the working of the group and attendance issues.

Activities:

- **Warm-up (each group member interviews another and introduces them to the group).**
- **A privacy contract was drawn up with the group members.**
- **A discussion about the small group study process and responsibilities within the group with the group members.**
- A discussion with the group regarding the day-to-day problems that they have and their relationship with their disease.
- Sharing communication problems with the group members.

2nd Session

Aim: To improve the perception of individuals with regard to their feelings.

Activities:

- Exercises and practice;
- The group discussed how they felt at the previous session.
- A demonstration about the definition of the emotion, the relation between emotion, thought and behavior, and basic emotions that people feel towards positive and negative situations.
- The group members listed the emotions they recalled in a poster-sheet and shared these with the group.
- All the lists were pinned to the wall and the most comprehensive lists of feelings were divided into negative (disliked) and positive (liked) examples. All these feelings were discussed with the group one by one, from the least frequent to the most common.
- Group discussion about the process.
- Homework involving identifying daily emotions

3rd and 4th Sessions

Aim: To differentiate between emotions.

Activities;

- Sharing group experiences regarding the homework from the 2nd session.
- A demonstration about the constitution of the emotions, the differences between basic and complicated emotions being aware or unaware of the emotions, the different features of emotions.
- Role-plays concerning either the group members' real-life experiences or scenario-based ones.
- The group practiced improving their body language regarding perceiving the emotions.
- The group members were given a list containing some 56 basic and complicated emotions. They were then asked to observe themselves and their relatives in terms of the emotions listed there in.
- Homework involving these exercises.

5th Session

Aim: To focus on awareness of methods by which the emotions are expressed.

Activities:

- The group shared their experiences of the homework from the 3rd and 4th sessions
- A demonstration of the personal differences in expressing emotions and their effect on human relations.
- The group practiced expressing basic and complicated emotions.
- A discussion with the group regarding how they express their feelings in their daily lives.
- Homework involving observing how close relatives or persons around express their feelings and sharing this with the group.

6th Session

Aim: To manage the emotions.

Activities:

- The group shared their experiences regarding the homework from the 5th session
- A demonstration of personal differences in the management of the emotions, where these differences may originate from, and the relation between emotional management and behavior.
- Group discussion of the topic
- The group practiced the management of emotions.
- The group members expressed their emotional management abilities using role-play regarding daily-life situations.
- Homework observing themselves, close relatives or friends involving emotional management.

7th and 8th Sessions

Aim: To improve empathic bonding abilities with others, displaying empathic reactions and attaining the skill of empathic listening.

Activities:

- The group shared their experiences regarding the homework from the 6th session
- A demonstration of empathy, the components thereof, and the process of bonding with empathic relations, and effective and ineffective empathic responses.
- The group practiced bonding and responding using advanced empathy skills.
- The group discussed whether or not their daily life responses were empathic.
- Homework to maintain group members' human relations using advanced empathic skills.

9th Session

Aim: To acquire self-motivational ability

Theme: To focus on learning to expend motivational energies toward a determined target and in a specific way, to use motivating and trusting speech.

Activities:

- The group shared their experiences regarding the homework from the 7th and 8th sessions.
- A demonstration of motivation, the resources (internal and external) of motivation, and activities increasing motivation.
- The group practiced increasing motivation.
- Role-plays including using motivating words to themselves in the face of scenario-based situations.
- Homework using motivation-increasing words in daily life.

10th and 11th Sessions

Aim: To use emotions in daily life

Theme: Focused on differentiating between forms of behavior, those which are friendly from those which are not, forming positive thoughts in friendships, being aware of the presence of multiple solutions to a given problem and developing skills for the management of relationships.

Activities:

- The group shared their experiences regarding the homework from the 9th session.
- A demonstration of the importance of controlling social relationships and the importance of emotions in order to achieve this goal.
- The group members practiced friendly and unfriendly words and forms of behavior.
- Role-plays using friendly and unfriendly words and behavior in daily life.
- Homework: using these abilities in daily life.

12th Session

Aim: To share the group's feelings regarding this program.

Theme: To evaluate the process regarding group life. The last session focused on sharing thoughts and emotions concerning the group experience, included a final evaluation of the program with the trainees and culminated with relaxation training.

Activities:

- The group members practiced relaxation techniques.
- Discussion of the program
- Feed-back

Bullying and School Climate from the Aspects of the Students and Teachers

Hülya Kartal*
Asude Bilgin**

Suggested Citation:

Kartal, H. & Bilgin, A. (2009). Bullying and school climate from the aspects of the students and teachers. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 209-226.

Abstract

Problem Statement: School culture and school climate describe the environment that affects the behavior of teachers and students. School climate reflects the physical and psychological aspects of the school that are more susceptible to change and that provide the preconditions necessary for teaching and learning to take place. In addition, school climate is a significant element in discussions about improving academic performance. The most important component of school climate is safety. However, incidents of bullying are frequent occurrences for many children at school and in the community. Children who are bullied fear coming to school, and they believe school to be an unsafe and distressing place.

Purpose of Study: The primary purpose of the present study is to investigate the feelings of both elementary students and their teachers about the school climate in terms of their perceptions of safety in the school. It is also of importance to describe the types and the prevalence of bullying, and possible incidents where bullying occurs, where they feel safe, and whom they tell about bullying.

Methods: 688 students from the 4th through 8th grade from one elementary school and 58 randomly selected teachers were presented a questionnaire called the Colorado School Climate Survey. The data was analyzed in terms of frequencies and variance.

Findings and Results: 41.3% reported that over the last month they were bullied at least once in a week or more. The results showed an average school climate scored 42.5 points out of a 56 point maximum. 25.4% of the students were found to be neither bullies nor victims. 41.3% of them were victims, 3.3% were bullies, and 29.9 % were bully-victims.

Conclusion and Recommendations: Bullying is a serious problem in many schools and there seems to be no one single solution for every student. It is known, however, that the social context and supervision of the school play a major role in the frequency and severity of bullying problems.

Keywords: School climate, bullying, elementary school students, teacher.

* Asst. Prof. Dr., Uludağ University Faculty of Education, hkartal@uludag.edu.tr

** Prof. Dr., Uludağ University Faculty of Education, asudebilgin@uludag.edu.tr

Arthur Perry (1908), at the beginning of the 20th century, was one of the first educationalists to clearly write about how school climate affects students and the process of learning (cited in Mcneely, Nonnemaker, & Blum, 2002). In the last decade, a growing body of educators and researchers has attempted to identify exactly what the definition of school culture and school climate is (Cohen, 2006). The terms school culture and school climate describe the environment that affects the behavior of teachers and students. School culture “reflects the shared ideas, assumptions, values, and beliefs that give an organization its identity and standard for expected behaviors.” School climate “reflects the physical and psychological aspects of the school that are more susceptible to change and that provide the preconditions necessary for teaching and learning to take place” (Tableman, 2004). School climate is also defined as the feelings that students and staff have about the school environment. These feelings may be related to the feelings of comfort, the perception about the supportiveness of learning or teaching environment, and the safety of the people in the school. It could be hypothesized that these kinds of feelings lead to effective and efficient learning and teaching as well as positive student behavior and attitudes (Peterson & Skiba, 2001).

School climate also refers to the quality and consistency of the interpersonal relations which take place among members of staff, staff and students, students and students, and staff and parents. In this context, school climate is thought to offer potential solutions to problems such as bullying, inter-student conflicts, suicide, and school absenteeism. These positive effects lead to fewer disciplinary incidents and more improved student academic performance (Haynes, Bennoss, & Ben-Avi, 1997). A safe, caring, participatory, and responsive school climate tends to foster great attachment to school (Osterman, 2000) and a connection to one another within the school ((Mcneely et al., 2002). Positive school climate also enhances teacher retention (Cohen, 2006).

Sometimes the school culture develops dysfunctional values and beliefs. Deal and Peterson (1998) call this dysfunction “toxic cultures”. In “toxic cultures”, the staff views students as the source of problem. They often complain, criticize, discourage, and demoralize. They do not share ideas or materials. In these cultures, staff is afraid to offer suggestions or new ideas for fear of being attacked or criticized. They tend to use verbal or physical punishments to correct the misbehavior of the students. In a negative school climate, students and teachers both dislike being in the school (Deal & Peterson, 1998 as cited in Peterson, 2002). For instance Rutter (1983) reported that in punitive schools, failure and absenteeism are strongly related to school climate. 86% of students who often miss school reported institutional reasons for doing so, such as bullying, school rules, punishments, and teachers (Reid, 1982). In a literature review, Tableman (2004) reports that a positive school climate is associated with higher grades, engagement, attendance, aspirations, on-time progression, higher self-esteem and self-concept, less anxiety, less depression and loneliness, and less substance use. A research conducted by the Yale Child Study Center School Development Program supported the hypothesis that a positive school climate is especially crucial for low SES students (Haynes et al., 1997).

Equipping the students with the skills needed to change the climate of the school is necessary to impact and lessen school violence (Garrity & Jens, 1997). It is almost

impossible to improve the quality of education without controlling school violence because in spite of the good the teachers or curriculum, violence spoils the school climate and of course the learning environment (Volkh & Snell, 1998). Bullying is the most prevalent form of violence (Orpinas & Horne, 2006). According to Farrington (1993) "bullying is repeated oppression of a less powerful person, physical or psychological, by a more powerful person."

Bully-victim problems take up the valuable school time of the staff. Some studies show that when individual members of school staff deal with bullies one-on-one there is high possibility for these individuals to fail. However when there is a school-wide agreement to end bullying, it can be reduced by up to 50%. In the previous years, very effective programs have been developed to reduce bullying in schools. An effective approach is "raising awareness about bullying, increasing teacher and parent involvement and supervision, forming clear rules and strong social norms against bullying, and providing support and protection for all students" (Olweus, Limber, & Mihalic, 1999; Fekkes, Pijpers, & Verloove-Vanhorick, 2005; Kanetsuna, Smith, & Morita, 2006; Egbochuku, 2007).

Garrity and Jens (1997) reported five key components for building a school culture with a safer school environment that does not tolerate acts of physical or emotional aggression by anyone:

Teachers and Other Staff

The intervention of the teachers is very important for the students. They need to know that the staff will intervene, but generally staff doesn't know how to intervene or they tell students to solve their problems on their own, or when they do intervene they use punishment which makes the situation worse. Many schools believe that punishment and power will make students safe. Whereas ending school violence is more difficult than trying to control the students (Randel, 2007).

The non bully or victims

Non-bullies or victims consist of the majority of children who a significant effect on the climate of the school. Most bullying happens when adults are not around, such as during recess, in the playground, and after school. Bullying rarely takes place without an audience such that 85 percent of the students witness bullying (National Crime Prevention Council, 2008). They don't like the bullying but they do not know how to intervene and cannot stop bullying. In cases when students do intervene, 57 percent of the bullying stops within 10 seconds (Hawkins, Pepler, & Craig, 2001).

Bullies

Bullies tend to be confident, with high self-esteem (Nansel et al., 2001). They are generally physically aggressive and usually have little empathy for their victims. Bullies come from homes where parents provide little emotional support for their children. Also, an extremely permissive or excessively harsh disciplinary approach can increase the risk of bullying (Olweus, 1978 as cited in Garrity & Jens, 1997). Bullies tend to channel their power and need for attention prosaically (Randel, 2007).

Victims

Victims are typically anxious, insecure, shy, and cautious and suffer from low self-esteem, rarely defending themselves or retaliating when confronted (Nansel et al., 2001; Arıçak, 2009). While most victims do not do anything to provoke the victimization, some may have a tendency to show irritating behaviors. These children tend to be impulsive and have poor social skills. These *provocative victims* may also try to bully other children (Olweus, 1993). Student observations have recorded that bullying happens every seven minutes and that an equal number of boys and girls report being victims (Craig & Pepler, 1997). Creating a safer school climate helps victims feel protected as they learn to develop better social skills.

Parents

Parents would like to know that their children are safe and cared for. The parent's responsibility is to know the difference between a bully-victim situation and normal conflict in their child's peer relations (Garrity & Jens, 1997). On the other hand, some disciplinary approaches of the parents can predict whether children will be aggressive bullies. Lack of interest, love, and empathy toward the child, together with modeling aggressive behavior at home and poor supervision of the child, creates the conditions for aggressive and bullying behavior to happen (Olweus, 1993).

The approach for bully-proofing the school involves teachers, principals, students, parents, and everyone associated with the school. The students need to be kept safe emotionally and socially, not just physically. Although there have been some publications on studies about bullying over the last decade, much work is still needed to be done in this field.

Purpose of Study

The main purpose of the present study is to determine the perceptions of the elementary students and the teachers about the school climate and bullying. The research questions were as follows: 1. To what extent are children involved in bullying behavior? 2. What are the types of bullying? 3. Where does the bullying take place? 4. To whom the bullying is reported? 5. How is the school climate perceived?

Method

Research model

This study was built up on a descriptive research model. The main goal of this type of research is to describe the data and characteristics about what is being studied. The idea behind this type of research is to study frequencies, averages, and other statistical calculations. Although this research is highly accurate, it does not aspire to find out the causes behind bullying.

Participants

Local Educational Authorities were applied to for permission to carry out a study in an elementary school in the city of Bursa. The permission was granted. The convenient classes which the school principal manager chose were used for the study. The participants were 688 students (349 girls, 334 boys, 5 gender anonymous)

from the 4th-8th grade and 58 teachers during the 2006-2007 academic year. The students were selected among the convenient students who were present and volunteered in the classes during the time of application. Table 1 displays the distribution of the participants according to the grades and gender:

Table 1
Gender and Grade Level of Elementary School Students

	Girl		Boy	
	Frequency	Percent	Frequency	Percent
Grade 4	107	15.6	116	17.0
Grade 5	75	11.0	55	8.1
Grade 6	50	7.3	52	7.6
Grade 7	68	10.0	64	9.4
Grade 8	49	7.1	47	6.9
Total	349	51.0	334	49.0

Instrument

The "Colorado School Climate Survey", which was developed as a part of a bully proofing program for schools by Garrity et al. (2000), was used for the study. The questionnaire was translated into Turkish and back translation was performed by two instructors from the English Language Teaching Department of Uludag University. The content validity was determined by the opinions of three educational psychologists with a PhD. The same questionnaire was used by Kartal and Bilgin (2007 and 2008). The questionnaire was designed to measure several aspects of school climate. Several subscales were used in the questionnaire: bullying experienced, bullying witnessed, the strategies they use when they are bullied or witness other students being bullied, student' perceptions of the school climate, who was bullying done by, where did it happen, who did the students tell, how safe do the students feel. The questionnaire was a 3- point likert type. The alpha coefficient of the internal consistency of the questionnaire was $\alpha = .69$ for the six items about types and prevalence of bullying, $\alpha = .59$ for five items about safety perceptions of participants in school, and $\alpha = .77$ for 14 items about school climate perceptions of the participants.

Procedure

Official permission from the National Education Management of the Bursa province was obtained before the application. The questionnaires were presented to the students during regular class hours. They were asked to reply to the questions honestly as a part of a research study. They were told not to write their names on the questionnaires. The questionnaire was administered by the researchers themselves.

Data Analysis

The data was analyzed using statistical technique of frequencies and an analysis of variance.

Findings and Results

The results of the study are presented below:

1. The Rates of Primary School Students Partaking in Bullying

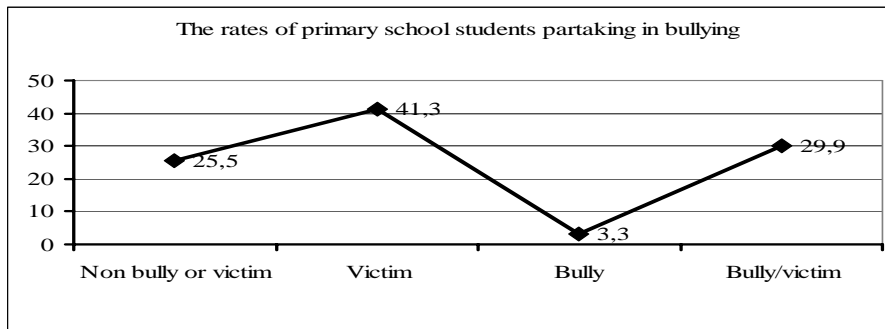


Figure 1

As shown in Figure 1, in the last month, 25.5 percent of elementary students were never included in bullying; 41.3 percent of them reported being bullied at least once a week or more; 3.3 percent of them admitted to have bullied others. Nearly 1/3 of them (29.9%) has bullied others and has been bullied. When these findings are considered as a whole, it can be concluded that the bullying is quite common in this school.

2. The Bullying Types and Prevalence According to the Students and Teachers

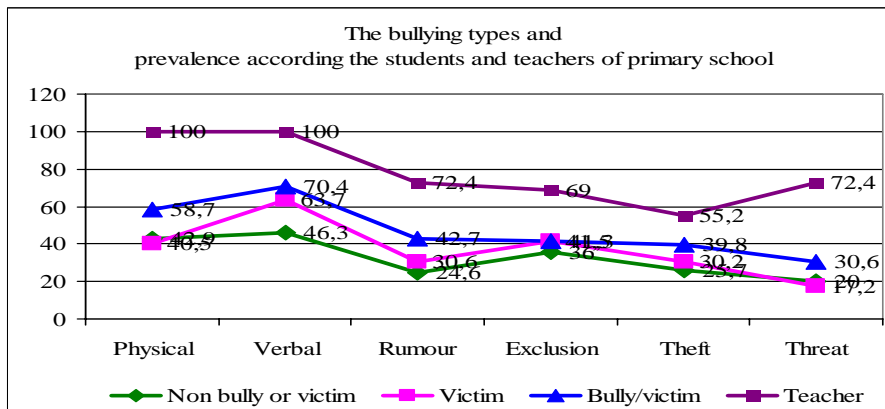


Figure 2

Figure 2 presents the bullying types and prevalence according to the students and teachers of the elementary school. The teachers rated all bullying types to be more prevalent than the students. At the highest rate were the bully/victims (70.4%), then victims (63.7%) and non bully victims (46.3%). Teachers reported that the most

frequent bullying type is verbal bullying in the form of saying mean things, teasing, and name calling.

3. According to the Reports of Students and Teachers Where Bullying Occurs

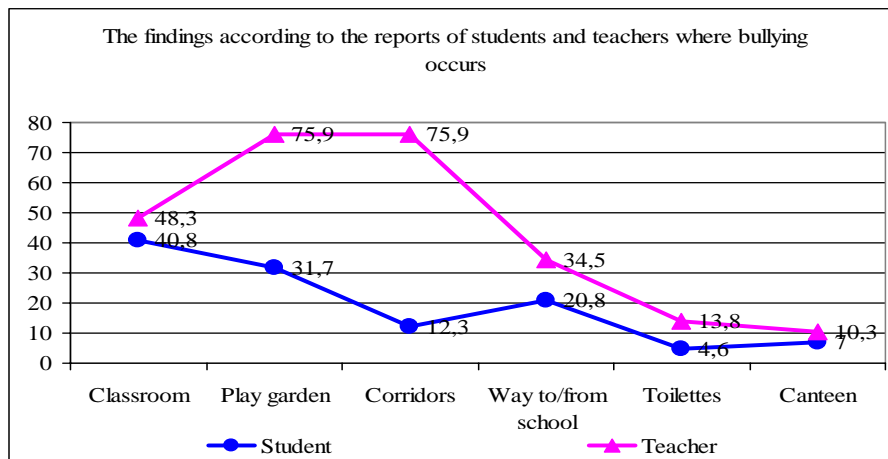


Figure 3

As can be seen from figure 3, it has concluded that bullying happened mostly in the classroom (40.8%). It is followed by the playground (31.7%), on the way to school (20.8%), hallways (12.3%), the canteen (10.3%), and then toilets (4.6 %). The majority of the teachers said that bullying happens mostly in the playground and hallways (75.9%). It was followed by the classroom (48.3%), on the way to/from school (34.5%), toilets (13.8%), and then the canteen (10.3%).

4. According to the Reports of Students and Teachers to Whom bullying is Reported

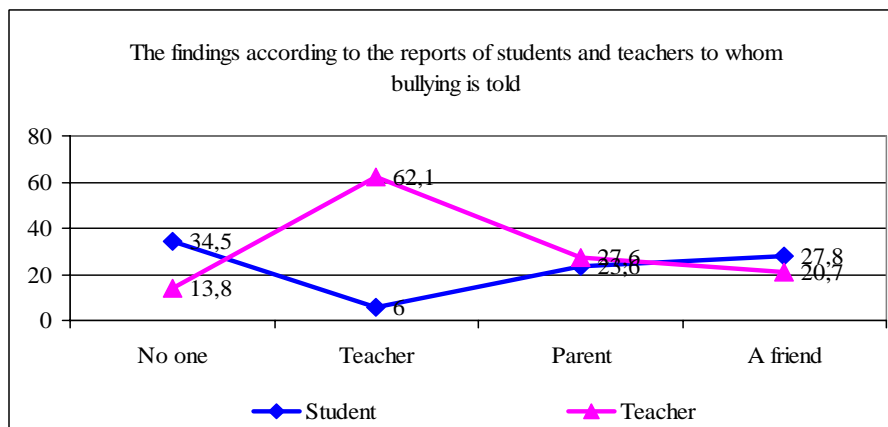


Figure 4

Figure 4 presents that these students who were bullied report this most frequently to their friends (27.8%) and then to parents (23.6%). Only 6.0% of them told to their teachers. 62.1 percent of teachers reported that students report their victimization to them (the teacher) first and then to friends (27.8%) and finally to parents (27.6%). 13.8 percent of the teachers believed that students do not tell anyone about their victimization, but 34.5 percent of the students reported the alternative.

5. Safety Perceptions of the Victims and Teachers

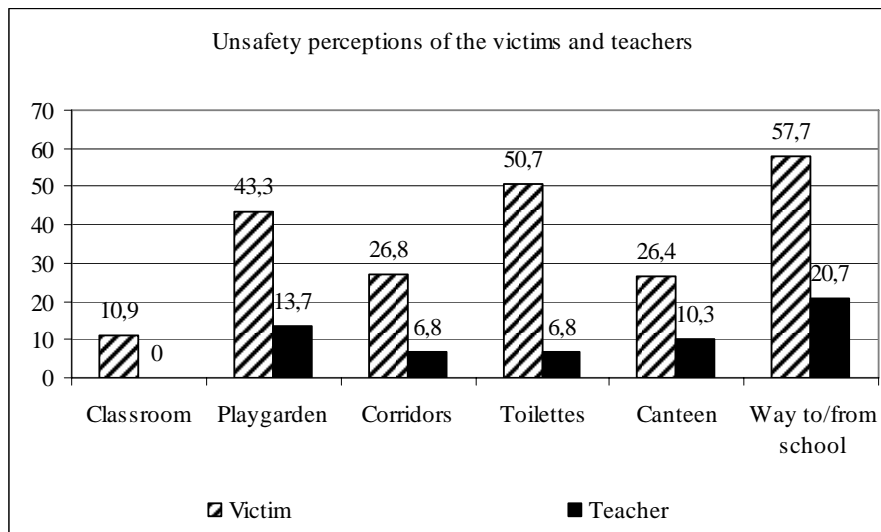


Figure 5

Figure 5 presents the safety perceptions of the victims and teachers. Only half of the students perceived the classrooms to be safe, but almost all the teachers reported them to be safe. In contrast to 43.3 percent of the students, only 13.7 percent of teachers reported the playground as unsafe. While half of the teachers (48.3%) perceive that hallways are safe, only 27.8% of the students reported them as safe.

6. School Climate Perceptions of Students and Teachers

Table 2

The School Climate Scores for the Students and Teachers

	N	Minimum	Maximum	Mean	Std. Dev.
Non bully or victim	175	17	56	43.9	7.1
Victim	284	18	56	42.3	7.1
Bully	23	34	53	43.8	5.6
Bully/victim	206	17	56	40.1	7.1
Teacher	58	30	53	41.2	5.9

From the analysis of the findings about school climate, it has seen that the mean scores of the teachers and students are very close to each other. The highest score

(43.9; total score = 56) was belonging to “the non bully or victim” group and the lowest (40.1; total score = 56) was to bully/victim group. A variance analysis was performed to investigate the significance of the differences among the school climate score of students and the teachers. The result of the variance analysis is shown in Table 3:

Table 3

The Analysis of Variance for Comparing School Climate Score Mean Levels of Participants

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	925.21	4	231.30	4.72	.001
Within Groups	31706.02	647	49.00		
Total	32631.24	651			

In table 3, the mean of the school climate scores of the groups were compared and the F value was found to be significant ($F=4.72$; $p<.05$). There is a significant difference between bully/victim and non bully or victim students. The school climate scores seem to change depending on the form of participation that the students took in the bullying process.

Discussion and Conclusion

This study was intended to investigate the perceptions of the elementary students and teachers about their school climate. In terms of this general purpose some research questions were investigated:

The first group of findings was related to the rates of elementary school students partaking in bullying. In this study, only one out of four of the students reported that they were not involved in any bullying incidences in the last one month. Nearly half of the children reported that they were bullied at least once or more in a week during the last month.

In general, these results seem to show that the prevalence of victimization and the rate of bully/victims is quite high in comparison with the findings of the related literature. According to recent studies, at least one out of three of all pupils has been a victim of bullying or has bullied classmates (Fox et al., 2003; Cohn & Canter, 2003). Yurtal and Cenkseven (2006) reported that nearly half of the students encountered bullying at some time. Erdur-Baker and Kavşut (2007) reported that 4.9%-30.1% of high school students engaged in cyber bullying. Fekkes and colleagues (2005) revealed that all students take place in bullying incidences somehow. In South Africa, Greeff (2004) found that more than half of the students were bullied at least once in the last term. Pişkin (2006) found that 35% of elementary students bullied steadily and Kapıcı (2004) declared that 40% of children from fourth and fifth grade have been exposed to bullying.

The teachers reported higher rates for all bullying types than the students did. There are similar results in the literature. For example Stockdale, Hangaduambo, Duys, Larons, and Sarvela (2002) found that the teachers reported the highest rates of physical and verbal bullying among students and parents. Boulton (1997) investigated the reports of teachers and students in preschool and elementary school students. The results showed that the teachers reported significantly higher bullying rates than the students. Also Naylor, Cowie, and Cossin (2006) found parallel results, but, on the other hand, in the study of Barone (1995), teachers reported lower bullying rates than the students (cited in Stockdale et al., 2002). It seems that the teachers are not aware of their responsibility to intervene and decrease bullying in their classes. Otherwise, in terms of social desirability, they probably would report lower rates.

In the current study, all of the teachers report that physical bullying is the most prevalent form of bullying. Verbal bullying is second. While teacher reports emphasize physical bullying, the student reports pointed out that verbal bullying was the most prevalent form of bullying. This result can be due to difficulties for teachers to perceive verbal bullying, because teachers are incapable of knowing it is bullying when they hear it. However in general, these findings are interpreted as an indication for needing to increase the teachers' awareness of bullying and the need for supporting the students in learning constructive conflict resolution strategies.

The students reported that bullying incidences took place mostly in the classrooms and the playground and seldom in the bathroom. The great majority of the teachers reported that bullying occurs usually on the playground and in the hallways, then in the classrooms. In general, in most of the studies, the playground was found to be the main place for bullying (Seals & Young, 2003; Çınkır & Karaman-Kepenekçi, 2003). Fekkes, Pijpers, and Verloove-Vanhorick (2005) found that bullying happens most in the order of hallways, classrooms, canteens, and then toilets. Dake and his colleagues (2003) declared that bullying mostly occurs on the playground and then in classrooms and hallways. In the study of Buchanan and Winzer (2001) students said that bullying happens on the playground because teachers are not there and even if they are present they do not see the incidences. In Turkey, Pişkin (2006) found that classrooms, canteens, and playground were the places where bullying frequently occurs. This is because playgrounds and classrooms are the two places where children mostly interact with each other suggest that more effective supervision is needed in these places. Olweus (1993) found that the level of bullying was lower in schools where there are relatively more teachers present during recess and lunch breaks.

Another significant result is the difference between the rates of teachers and students reporting bullying incidences in bathrooms. The teachers rated the bathroom as a place of incidence three times more than students did. Bathroom bullying may include vandalism, smoking, drinking, drug use, and sex. Byles (2007) listed some suggestions for a solution to bathroom bullying (Praeger, 2007): (1) "Making hand-washing areas unisex and more visible; (2) Locating bathrooms

opposite to classrooms so they can be more easily supervised; (3) Employing a full-time restroom attendant; (4) Playing music; (5) In extreme cases, monitoring students, cameras are best used when other options have failed; And (6) dead-end corridors and dark corners should be avoided and the general visibility in all parts of schools will be increased.

According to the reports of students, only a few of them tell their teachers that they are bullied, but a great majority of the teachers said that victims tell them about their situation. A significant difference was found between these rates. In fact, the students stated that they were mostly telling their friends and then their families when they were bullied. Even the "telling no one" alternative was reported twice as much as the alternative, "I told an adult in the school" (Kartal & Bilgin, 2007). Houndumadi and Pateraki (2001) found parallel results. In England, 51.3 percent of second grade students and 35.7 percent of fourth grade students said that they were always telling the situation to their teachers. In Germany, this rate is just 9.8 percent (Wolke et al., 2001a). A high proportion of English pupils mentioned that victims may be afraid of the bullying getting worse if they tell. This points to the importance of teachers responding sensitively but firmly and effectively when a victim does seek help (Kanetsuna et al., 2006). Students do not tell their teachers because they do not believe that they will intervene successfully. In the study of Kepenekçi and Çınkır (2006), students reported that they did not get any help from adults. When all the findings are taken as a whole, it seems clear that most of the students do not tell their teachers and parents or get help when they are bullied.

However, in a school which has a good climate the students are expected to report bullying to their teachers. Otherwise it can be concluded that the interaction between students and teachers is not sufficient. Encouraging other children to intervene when they notice bullying behavior can also be an important strategy to combat bullying. The intervening of bystanders to help the victim is known to be effective. Motivating children to intervene and stand up for the victim could isolate bullies from their audience and their social support and help to stop bullying behavior (Fekkes et al., 2005). According to Kanetsuna and his colleagues (2006), an important step to prevent bullying is to change the climate of the school so that victims of bullying can tell others about being bullied with trust and confidence. Also, peers as well as teachers who witness bullying should feel empowered to act helpfully.

Findings about the safety perceptions of the victims and teachers revealed that both teachers and victims think the classroom as the safest place. Nearly half of the students versus a third of the teachers evaluate the playground as safe. These results are parallel to the results of the "where bullying occurs" question. Half of the students declared toilets as unsafe, yet only a small percentage of students reported that bullying takes place in toilets. This may be due to the cultural acceptance of the issues involved with toilets and probably also due to students feeling ashamed to talk about bullying incidences in toilets. This is very important, because sexual bullying can take place around these areas. That is why there is a great necessity for supervision of washrooms.

It is supposed to be another important finding that more than half of the students and a fifth of the teachers reported on the way to/from school as unsafe. This is not similar to the findings of the amount of bullying reports related to this area. Various authors have noted that there is much more bullying in school than there is on the way to and from school (Olweus, 1978; Ziegler & Rosenstein-Manner, 1991; Rivers & Smith, 1994 cited in Bidwell, 1997). Generally schools claim that bullying on the way to or from school is not their responsibility but teachers have to advise children to walk home with a group of friends or get an adult to supervise them home.

The last findings were about the school climate scores of the students and teachers. This study focused on bullying as an important factor of school climate. Results revealed that quite a large amount of the students were bullied at least once or more a week in the last month. Nearly 1/3 of them reported that they bullied somebody and were bullied at some time. These results clarify the fact that bullying is very common in this elementary school. The highest school climate score belongs to non bullies or victims. The lowest score belongs to the bully/victim group. A significant difference was found between the scores of these two groups. This means that the bully/victims are the most negatively effected group over all.

There are two important factors that allow bullying to continue to prosper. The first is the non bullies or victims who avoid intervening and the second is the authorities who are ignoring the problem. The research presents us a general and dramatic picture: In average, 50 percent of students are bullied at some time, 10 percent regularly. This means that every seven minutes a child is bullied. Out of these situations 4 percent of the time adults intervene, 11 percent of the time peers intervene, and 85 percent of the time no intervention occurs (Randel, 2007).

Bullying is a serious problem facing many schools and of course there is no one single solution for every student. However, it has to be known that the social context and supervision at schools has been playing a major role in the frequency and severity of bullying problems. Bullying prevention efforts can increase feelings of safety for everyone in the school and help students learn to behave in nonviolent ways. Providing adult supervision during recess and in the hallways and toilets, staying in contact with parents, training school personnel to detect and intervene, and dealing with the situation with fairly consistent discipline methods are all importance factors for overcoming this problem (Novince, 2007). As Randel (2007) said "while teachers and administrators do not have control over individual and family factors which produce children who are inclined to bully, bullying problems can be greatly reduced in severity by appropriate supervision, intervention and climate in a school." A positive school climate exists when all students feel comfortable, wanted, valued, accepted, and secure in an environment. Improved school climate is a goal to pursue everywhere. Educators need to constantly work toward improving their school climate, culture, and conditions so that student learning is improved. This is not a problem that can be completely solved overnight, but it can be controlled. If parents, teachers, administrators, school staff, and the neighborhood community acknowledge it and each does their part, then it can be reduced and eventually prevented entirely (Randel, 2007).

References

- Ancak, O.T. (2009). Psychiatric symptomatology as a predictor of cyberbullying among University students (Üniversite öğrencilerindeki siber zorbalık davranışlarının bir yordayıcısı olarak psikiyatrik belirtiler). *Eurasian Journal of Educational Research*, 34, Winter, 167-184.
- Batsche, G.M. & Knoff, H.M. (1994). Bullies and their victims: Understanding a pervasive problem in the schools. *School Psychology Review*, 22, 165-174.
- Bidwell, N.M. (1997). The nature and prevalence of bullying in elementary schools. A summary of a master's thesis. Retrieved October 16 2007 from http://saskschoolboards.ca/research/school_improvement/97-06.htm
- Boulton, M. J. (1997). Teachers' views on bullying: Definitions, attitudes and ability to cope. *British Journal of Educational Psychology*, 67, 223-233.
- Buchanan, P., & Winzer, M. (2001). Bullying in schools: Children's voices. *International Journal of Special Education*, 16(1), 67-79.
- Bureau of Justice Statistics and National Center for Education Statistics, Indicators of School Crime and Safety Statistics. (2005). Retrieved September 6 2007 from <http://www.insurancejournal.com/magazines/midwest/2006.12.11/features/75333.htm>
- Cohen, J. (2006). Social, emotional, ethical, and academic education: Creating a climate for learning, participation in democracy and well, being. *Harvard Educational Review*, 76(2), 201-237.
- Cohn, A. & Canter, A. (2003). Bullying: Facts for Schools and Parents. *National Association of School Psychologists*. Retrieved from <http://www.naspcenter.org>.
- Craig, W. M. & Pepler, D. J. (1997). Observations of bullying and victimization in the school yard. *Canadian Journal of School Psychology*, 13(2), 41-60.
- Dake, J.A., Price, J.H., Telljohann, S.K., & Funk, J.B. (2003). Teachers' perceptions and practices regarding school bullying prevention activities. *Journal of School Health*, 73, 347-354.
- Egbochuku, E.O. (2007). Bullying in Nigerian schools: Prevalence study and implications for counseling. *Journal of Social Sciences*, 14(1):65-71. Retrieved December 26 2007 <http://www.krepublishers.com/02-Journals/JSS/JSS-14-0-000-000-2007-Web/JSS-14-0-000-000-2007-Contents/JSS-14-0-000-000-2007-Contents.htm>
- Erdur-Baker, Ö. & Kavşut, F. (2007). Cyber bullying: A new face of peer bullying (Akran zorbalığının yeni yüzü: Siber Zorbalık). *Eurasian Journal of Educational Research*, 27, Spring, 31-42.
- Farrington, D.P. (1993). Understanding and Preventing Bullying. In M. Tonry (Eds.), *Crime and Justice. A Review of Research*, 17, 381-458.
- Fekkes, M., Pijpers, F. I. M., Verloove-Vanhorick, S.P. (2005). Bullying: Who does what, when and where? Involvement of children, teachers and parents in bullying behavior. *Health Education Research*, 20(1), 81-91.

- Fox, J.A., Elliott, D.S., Kerlikowske, R.C., Newman, S.A. & Christeson, W. (2003). Bullying prevention is crime prevention: A report by Fight Crime: Invest in kids. Retrieved March 30 2008 from <http://www.fightcrime.org/reports/BullyingReport.pdf>
- Garrity, C., & Jens, K. (1997). Bully proofing your school: Creating a positive climate. *Intervention in School & Clinic*, 32(4), 235-244.
- Garrity, C., Jens, K., Porter, W., Sager, N., & Short-Camilli, C. (2000). *Bully-proofing your school: A comprehensive approach for elementary school*. Chapter Three: Colorado school climate survey, elementary school student report, (pp.55-61) (2nd Edi.), Longmont, CO: Sopris West.
- Greeff, P. (2004). *The nature and prevalence of bullying during the intermediate school phase*. Unpublished thesis of master, Faculty of Humanities, Department of Psychology, University of the Free State. Retrieved June 11 2007 from <http://etd.uovs.ac.za>
- Hawkins, D.L., Pepler, D.J., & Craig, W.M. (2001). Naturalistic observations of peer interventions in bullying. *Social Development*, 10 (4), 512-527.
- Haynes, N.M., Emmons, C., & Ben-Avie, M. (1997). School climate as a factor in student adjustment and achievement. *Journal of Educational and Psychological Consultation*, 8 (3), 321-329.
- Houndoumadi, A. & Pateraki, L. (2001). Bullying and bullies in Greek elementary schools: pupils' attitudes and teachers' /parents' awareness. *Educational Review*, 53(1), 19-26.
- Humphreys, J. (1997). Nursing care of children of battered women. *Pediatric Nursing*, 23(2), 1-13.
- Kanetsuna, T., Smith, P.K., & Morita, Y. (2006). Coping with bullying at school: children's recommended strategies and attitudes to school-based interventions in England and Japan. *Aggressive Behavior*, 32, 570-580.
- Kartal, H. & Bilgin, A. (2007). İlköğretim öğrencilerine yönelik bir zorbalık karşıtı program uygulaması: Okulu zorbalıktan ayrılma programı [An implementation of an anti-bullying program for elementary school students: Bully proofing your school]. *Eğitimde Kuram ve Uygulama*, 3(2), 207-227.
- Kartal, H. & Bilgin, A. (2008). Öğrenci, veli ve öğretmen gözüyle ilköğretim okullarında yaşanan zorbalık. [Bullying in the elementary schools: from the aspects of the students, the teachers and the parents]. *İlköğretim Online*, 7(2), 485-495.
- Kepenekçi, Y. & Çınkır, S. (2006). Bullying among Turkish high school students. *Child Abuse & Neglect*, 30(2), 193-204.
- Kumpulainen, K., Rasanen, E., & Henttonen, I. (1999). Children involved in bullying: Psychosocial disturbance and the persistence of the involvement. *Child Abuse & Neglect*, 23(12), 1253-1262.
- McNeely, C.A., Nonnemaker, J.M., & Blum, R.W. (2002). Promoting student connectedness to school: Evidence from the National Longitudinal Study of Adolescent Health. *Journal of School Health*, 72 (4), 138-146.
- Nansel, T.R., Overpeck, M., Pilla, R.S., Ruan, W.J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association*, 285(16), 2094-2100.

- National Crime Prevention Council. (2008). Retrieved February 3 2008 from <http://www.ncpc.org/topics/by-audience/parents/bullying/non-bully-or-victims/>
- Naylor, P., Cowie, H., & Cossin, F., de Bettencourt, R., & Lemme, F. (2006). Teachers' and pupils' definitions of bullying. *British Journal of Educational Psychology*, 76(3), 553-576.
- Novince, L. C. (2007). Anti-Bullying Policies: Promoting a Positive School Climate Based on the State Board of Education's Anti-Harassment, Anti-Intimidation and Anti-Bullying Model Policy. Retrieved March 26 2008 from <http://www.ode.state.oh.us/GD/DocumentManagement/DocumentDownload.aspx?DocumentID=41553>
- Olweus, D. (1993). *Bullying at school: What we know and what we can do*. Cambridge, MA: Blackwell Publishers.
- Olweus, D., Limber, S., & Mihalic, S. (1999). *Blueprints for violence prevention*, Book Nine: Bullying prevention program. Boulder, CO: Center for the Study and Prevention of Violence.
- Orpinas, P. & Horne, A. M. (2006). Bullying prevention: Creating a positive school climate and developing social competence. (Brief article)(Book review). *Adolescence*, 41(163), 591-592.
- Osterman, K.F. (2000). Students' need for belonging in the school community. *Review of Educational Research*, 70, 323-367.
- Pepler, D.J., Craig, W.M., Ziegler, S. & Charach, A. (1994). An evaluation of an anti-bullying intervention in Toronto schools. *Canadian Journal of Community Mental Health*, 13(2), 95-110.
- Peterson, K.D. (2002). Positive or negative? *Journal of Staff Development*, 23(3), 10-15.
- Peterson, R. L. & Skiba, R. (2001). Creating school climates that prevent school violence. *Social Studies*, 92(4), 167-176.
- Pişkin, M. (2006, Mart). *Akran zorbalığı olgusunun ilköğretim öğrencileri arasındaki yaygınlığının incelenmesi* [Investigation of prevalence of peer bullying among elementary students]. I. Şiddet ve Okul: Okul ve Çevresinde Çocuğa Yönelik Şiddet ve Alınabilecek Tedbirler Sempozyumu'nda sunulan bildiri, İstanbul.
- Praeger, D. (2007). No behind left behind: can school bathrooms be fixed? Retrieved April 5 2007 from http://www.poopreport.com/BMnewswire/no_behind_left_behind.html
- Randel, D. (2007). Stopping school violence: A Bullying Program for Schools. Retrieved April 17 2008 from http://www.stoppingschoolviolence.com/articles/InSchool/articles_School_bullyPrg.html
- Reid, K. (1982). Retrospection and persistent school absenteeism. *Educational Research*, 25, 110-115.
- Rigby, K. & Slee, P.T. (1991) Bullying among Australian school children: reported behaviour and attitudes to victims. *Journal of Social Psychology*, 131, 615-627.
- Rutter, M. (1983). School effects on pupil progress: Research findings and policy implications. *Child Development*, 54, 1-29.
- Seals, D. & Young, J. (2003). Bullying and victimization: Prevalence and relationship to gender, grade level, ethnicity, self-esteem, and depression. *Adolescence*, 38, 735-747.

- Smith, P.K. & Shu, S. (2000). What good schools can do about bullying: findings from a survey in English schools after a decade of research and action. *Childhood*, 7(2), 193-212.
- Smith, P.K. (2008). Definition, types and prevalence of school bullying and violence. Unit for School and Family Studies. Goldsmiths College, University of London. Retrieved March 16 2008, from <http://www.oecd.org/dataoecd/27/47/33866548.ppt>
- Stockdale, M. S., Hangaduambo, S., Duys, D., Larons, K., & Sarvela, P. D. (2002). Rural elementary students', parents', and teachers' perceptions of bullying. *American Journal of Health Behavior*, 26, 266-277.
- Tableman, B. (2004). School climate and learning. Best practices brief. No: 31. A product of the University-Community Partnerships at Michigan State University. Retrieved January 24 2008 from <http://outreach.msu.edu/bpbriefs/issues/brief31.pdf>
- Volokh, A. & Snell, L. (1998). School Violence Prevention: Strategies to Keep Schools Safe (Unabridged). Policy Study No. 234. Retrieved December 18 2007 from <http://www.reason.org/ps234.html>
- Yurtal, F. & Cenkseven, F. (2006, Mart). *İlköğretim okullarında zorbalığın incelenmesi*. [Investigation of bullying in elementary schools] I. Şiddet ve Okul: Okul ve Çevresinde Çocuğa Yönelik Şiddet ve Alınabilecek Tedbirler Sempozyumu'nda sunulan bildiri, İstanbul.
- Wolke, D., Woods, S., Stanford, K., & Schulz, H. (2001). Bullying and victimization of primary school children in England and Germany: Prevalence and school factors. *British Journal of Psychology*, 92, 673-696.
- Zindi, F. (1994). Bullying at boarding school: a Zimbabwe study. *Research in Education*, 51, 23-32.

Öğrenci ve Öğretmenlerin Gözüyle Zorbalık ve Okul İklimi

(Özet)

Problem Durumu: Okul iklimi öğrenci ve okul çalışanlarının okul ortamı için besledikleri duygular olarak tanımlanmaktadır. Bu duygular rahatlık, güvenlik ve öğrenme ve öğretme ortamının destekleyiciliğine yönelik algılarla bağlantılıdır. Okul iklimi okul içindeki etkileşimin kalite ve tutarlığını yansıtmaktadır. Bu etkileşim okul elemanları, okul elemanları ve öğrenciler, öğrenciler ve öğrenciler, okul elemanları ve veliler arasında yaşanmaktadır. Bu bağlamda okul ikliminin zorbalık, çatışma, intihar ve devamsızlık gibi problemler için potansiyel çözümler üretebileceği kabul edilmektedir. Sonuçta bu olumlu etkiler okuldaki disiplin sorunlarının azalmasına ve öğrenci başarısının artmasına yardımcı olacaktır. Olumlu bir okul iklimi öğretmenlerin de okula bağlanmasını sağlayacaktır. Öğrencilerin okul iklimini olumlu yönde değiştirecek biçimde eğitilmesi okulda yaşanan şiddette de büyük bir azalmaya yol açacaktır. Eğitimin kalitesini geliştirmek okuldaki şiddet kontrol altına

alınmadığı sürece olanaksızdır. Zorbalık ise okulda yaşanan şiddetin en yaygın şeklidir.

Araştırmanın Amacı: Okulu zorbalıktan arındıracak tüm yaklaşımlar öğretmen, yönetici, öğrenci, aile ve okulla ilgili tüm kişilerin katılımını gerektirmektedir. Öğrenciler sadece fiziksel olarak değil duygusal ve toplumsal olarak da güvenlik içinde korunmalıdır. Zorbalık son yıllarda üzerinde durulan bir çalışma alanı olmasına karşın henüz cevaplanılacak pek çok soru bulunmaktadır. Bu çalışmanın temel amacı ilköğretim okullarındaki öğretmen ve öğrencilerin okul iklimi algılarını güvenlik duyguları bağlamında incelemektir. Bu genel amaç doğrultusunda öğretmen ve öğrenciler açısından okulda yaşanan zorbalığın tipi ve sıklığı, zorbalığın yaşandığı yerler, güvende hissedilen yerler, zorbalığın anlatıldığı kişiler betimsel bir yaklaşımla ele alınmıştır.

*Araştırmanın Yöntemi:*Bursa'da bir ilköğretim okulunun 4-8.sınıf öğrencileri arasından rastgele seçilen 688 öğrenci (349 kız 334 erkek) ve 58 öğretmen çalışmanın katılımcılarını oluşturmuşlardır. Çalışmada ölçüm aracı olarak zorbalık karşıtı bir program uygulaması için Colorado Okul İklimi Bilgi Toplama Formu kullanılmıştır. Araç okul ikliminin çeşitli yönlerini ölçmek için geliştirilmiş çeşitli alt ölçekler içermektedir: Yaşanan zorbalık, tanık olunan zorbalık, zorbalıkla başa çıkmak için kullanılan taktikler, okul iklimi algıları, zorbalığın kim/ler tarafından yapıldığı, nerede olduğu ve güvenli hissedilen yerler. Araç 3 seçenekli likert tipi sorulardan oluşmaktadır ve iç tutarlığı zorbalığın tipi ve sıklığı ile ilgili alt ölçekteki altı madde için $\alpha = .69$, katılımcıların güvenlik algılarına yönelik beş madde için $\alpha = .59$ ve okul iklimi algılarına yönelik 14 madde için $\alpha = .77$ olarak hesaplanmıştır. Uygulama 2006-2007 eğitim-öğretim yılının bahar yarıyılında, öğrencilere sınıflarında araştırmacılar tarafından yapılmıştır. Öğrenciler, uygulamanın başında araştırma konusu hakkında bilgilendirilmiştir. Öğrencilerden okul adlarını, sınıf düzeylerini ve cinsiyetlerini yazmalarını istenmiştir.

*Araştırmanın Bulguları:*Bulguların birinci grubu öğrencilerin zorbalıkta yer alma oranları ile ilgilidir. Öğrencilerin %25.5'i son bir ay içinde hiçbir zorbalık vakasında yer almadığını bildirirken %34'ü haftada en az bir kere olmak üzere zorbalığa uğradığını bildirmiştir., Bu bulgu ilgili yazın göz önüne alındığında oldukça yüksek görünmektedir. Öğretmenlere göre tüm zorbalık tipleri öğrencilerin rapor ettiği oranlardan daha yüksek bulunmuştur. Genel olarak bu bulgunun karşıtını elde etmiş araştırma sonuçlarının varlığına karşın benzer çalışma sonuçları da mevcuttur. Öğretmenlerin tümüne göre fiziksel ve sözel zorbalık haftada en az bir kez gerçekleşmektedir. Ancak tüm öğretmenler fiziksel zorbalığı en sık yaşanan zorbalık olarak rapor ederken, öğrenciler sözel zorbalığı birinci sıraya koymaktadır (%70.4-46.3 arasında değişen oranlarda). Bu bulgular öğretmenlerin sözel zorbalığı algılamada güçlük çektiği biçiminde yorumlanmakta ve öğretmenlerin bu konuda eğitilmesi gerektiğine dikkat çekmektedir. Öte yandan öğrenciler zorbalık vakalarının çok sınıfta, daha sonra bahçede gerçekleştiğini rapor etmişlerdir. Tuvaletler en son sırada yer almaktadır. Öğretmenlerin büyük çoğunluğu ise zorbalığın en sık yaşandığı yerler olarak sırasıyla bahçe, koridor ve sınıfları saymaktadırlar. Tüm bu alanlar aslında öğretmenlerin gözetim ve denetiminde olması beklenen yerlerdir.

Araştırmanın Sonuçları ve Önerileri: Sonuçlar genel olarak benzer araştırma sonuçlarıyla tutarlı değildir. Okul bahçesi öğrenciler tarafından araştırmaların çoğunda zorbalığın en sık yaşandığı alan olarak tanımlanmaktadır. Öğrenciler buna neden olarak öğretmenlerin bahçede olmamalarını, olsalar da zorbalığı fark etmemelerini göstermişlerdir. Öğrencilerin sadece %6'sı zorbalığa uğradığını öğretmenlerine anlattığını söylerken öğretmenlerin %62.1'i öğrencilerin bu durumu ilk önce kendilerine anlattığını ifade etmiştir. Aradaki fark dikkat çekicidir ve okul ikliminin olumlu olduğu bir okulda öğrencilerin sorunlarını öncelikle okuldaki yetişkinlere anlatması beklenmektedir. Güvenlik algıları söz konusu olduğunda hem öğretmenler hem de öğrenciler sınıfı en güvenli yer olarak rapor etmiştir. Öğrencilerin yarısı, öğretmenlerin de 1/3'ü bahçeyi güvenli bir yer olarak tanımlamıştır. Bu bulgu daha sık zorbalık yaşansa da öğrencilerin sınıfta kendilerini daha güvenli hissettiklerini göstermektedir. Bir diğer önemli bulgu öğrencilerin yarıdan fazlasının (%57.7) ve öğretmenlerin 1/5'inin (%20.7) okul yolunu güvensiz olarak değerlendirmesidir.

Son bulgu katılımcıların okul iklimi puanlarıyla ilgilidir. En yüksek okul iklimi puan ortalaması ne zorba ne de kurban olan gruba, en düşük puan ortalaması ise zorba/kurban olanlara aittir. Zorbalığın önlenememesine iki neden gösterilmektedir. Birincisi ne zorba ne de kurban olanların zorbalık durumunda müdahale etmeye çekinmeleri, ikincisi de okuldaki yetişkinlerin bu konuyu ciddiye almamalarıdır. Oysa zorbalık okullarda çok ciddi bir problemdir ve bu sorunun tek bir çözümü yoktur. Ancak okuldaki olumlu sosyal ortam ve sağlıklı denetim bu problemin önlenmesinde ilk ve en güçlü adımları oluşturacaktır. Bu nedenle eğitimciler okullarının iklimini ve kültürünü geliştirme ve iyileştirme için sürekli çaba göstermek zorundadır.

Anahtar kelimeler: Okul iklimi, zorbalık, ilköğretim öğrencileri, öğretmen

Development of a Two-Tier Diagnostic Test to Determine Students' Understanding of Concepts in Genetics

Didem Kılıç*
Necdet Sağlam**

Suggested Citation:

Kılıç, D., & Sağlam, N. (2009). Development of a two-tier diagnostic test to determine students' understanding of concepts in genetics. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 36, 227-244.

Abstract

Problem Statement: The effectiveness of science instruction, which aims to assure students' complete and accurate understanding of science concepts, can only be effectively evaluated when it is measured using the appropriate measurement tools. Concept maps, multiple-choice tests, and interviews - all of which are employed in determining students' understanding of science concepts - have been shown to be successful methods. However, each has its own limitations. In determining students' understanding of concepts, the use of two-tier tests - which evaluate students' understanding more effectively and which are easy to score and apply - is recommended.

Purpose of Study: One of the topics of biology that students experience great difficulty in understanding is genetics. This study aims at developing a two-tier test which is able to effectively evaluate secondary education students' understanding of the concepts of genetics, and conducting the validity and reliability analyses of the test.

Methods: The two-tier test, which is aimed at determining secondary education students' understanding of the fundamental concepts of genetics, was developed using a three-phase method. In the first phase, the test content was defined. In the second phase, the relevant literature was examined so as to determine students' misunderstandings; students were interviewed, and they were given a multiple choice test with free-response answers in which they had to explain the reasons for their answers. In the third phase, 14 two-tier test items which were composed of multiple choice content questions in the first tier and a set of five justifications for every potentially selected response to the first tier were developed. The test was applied to 231 students, and then item analysis was performed. Test reliability was calculated using the Kuder-Richardson 20 formula. In order

* Research Assistant, Hacettepe University Faculty of Education, didemk@hacettepe.edu.tr

** Prof. Dr., Hacettepe University Faculty of Education, saglam@hacettepe.edu.tr

This paper is an extended version of a paper presented at the World Conference on Educational Sciences (WCES2009), North Cyprus, 2009.

to determine construct validity, the method of group differences was employed.

Findings and Results: The discrimination indices of items were found to be between 0.34 and 0.74. Item difficulties fell in the range of 0.25 to 0.67, and the test difficulty was found to be 0.43. Test reliability was calculated as 0.86.

Conclusions: All the findings obtained from the item analysis, in addition to the findings from the validity and the reliability studies; demonstrate that the test may be utilized to determine secondary education students' understanding of the concepts of genetics concepts in a valid and reliable way.

Recommendations: Besides making a more effective evaluation possible, two-tier tests facilitate application and scoring. Therefore, we believe that developing and using such two-tier diagnostic tests in various fields will make significant contributions to science instruction.

Keywords: Two-tier diagnostic test, students' understanding, genetics concepts, misunderstanding.

One of the questions that must be answered in order for science education programmes to attain their goals is to what extent students understand science concepts and topics. In order to evaluate students' understanding of the concepts, the knowledge held by students and the foundation on which they base their knowledge must be determined. The effectiveness of science instruction, which aims to assure students' complete and accurate understanding of scientific concepts, can only be accurately evaluated when it is measured using the appropriate measurement tools.

Concept maps, multiple-choice tests, and interviews, all of which are employed in determining students' understanding of scientific concepts, have been shown to be successful methods. However, each has its own limitations. It takes time for both teachers and students to master concept maps; and their analysis requires additional instruction. Although interviews specify students' understanding of science topics, they are difficult to apply in class; since structuring, transcribing, and interpreting them requires time and experience (Odom & Barrow, 1995). There is a body of research which shows that multiple choice tests are the second most commonly used method, followed by interviews, in determining students' understanding of concepts (Palmer, 1998). Multiple-choice tests, a method which can be employed by teachers easily in class, determine students' knowledge of the topic; but, they fail to evaluate the rationales on which the knowledge is based (Odom & Barrow, 1995; Peterson, Treagust, & Garnett, 1989; Tan, 2000). Tamir (1989) supports the view that multiple-choice tests have certain limitations, and recommends that the selection of multiple-choice test items should contain the known alternative conceptions, and that students should give a reason for their answers.

Haslam and Treagust (1987), Peterson et al. (1989), and Treagust (1988) suggest using two-tier multiple-choice test items which contain alternative concepts that are based on the rationales behind student answers. The items in two-tier multiple-choice diagnostic instruments are specifically designed to identify alternative conceptions and misunderstandings in a limited and clearly defined content area

(Tan, Goh, Chia, & Treagust, 2002). According to Wang (2004), two-tier tests represent a more developed approach than previous methods as they consider students' rationales and interpretations for their answers. Moreover, it has been stated that, compared with other approaches, two-tier tests are valid and reliable instruments which evaluate student ideas more effectively, and which are easy to score and apply (Chen, Lin, & Lin 2002; Tan et al., 2002; Treagust & Chandrasegaran, 2007).

The first tier of items available in two-tier tests is usually composed of a content knowledge question containing two or three choices. The second tier of each item is made up of the possible reasons for the choices that are available in the first tier. These reasons, which form the choices of the second tier, are obtained from previously asked free-response multiple choice tests, interviews, and literature (Treagust, 2006). The requirement of choosing a reason in two-tier tests is a sensitive and effective way to measure students' meaningful learning (Tamir, 1989).

A wide range of two-tier tests specific to various content fields have been developed and used to determine students' understanding of science concepts (Treagust, 2006; Treagust & Chandrasegaran, 2007). One of the features common to these tests is that they were developed in relation to topics which were difficult and confusing for students, and which had the most alternative conceptions. One of the biology topics that students have the most difficulty in understanding is "genetics". Research studies have demonstrated that most teachers and students hold the view that genetics topics are very difficult both to learn and to teach (Bahar, Johnstone, & Hansell, 1999; Rotbain, Marbach-Ad, & Stavy, 2005; Tekkaya, Özkan, Sungur, & Uzuntiryaki, 2001).

Genetics is a fundamental part of biology and is also directly related with daily life (Knippels, Waarlo, & Boersma, 2005). In their study, Lewis, Leech, and Wood-Robinson (2000a; 2000b; 2000c), attracted attention to the increasing political, economic, ethical, and educational importance of genetics education. It is emphasized in their studies that new improvements and research studies are increasingly becoming the subject matter of news reports and that societies should be literate in genetics. In other words, they point out that individuals who graduate with a secondary education should be able to understand what they read or hear about genetics, and should be able to participate in the public debate.

Research studies into determining students' understanding of the concepts of genetics show that students' knowledge of these concepts, their functions, and their relations, is incomplete and confused. Tekkaya et al. (2001); in addition to stating that students confuse such concepts as chromosomes, genes, alleles, chromatids, and DNA; reveal that students have difficulty in understanding topics which reference genes and chromosomes, while also struggling with the study of genetics for which these concepts form a basis. Moreover, it was also found that students could not accurately solve problems because they did not understand the relationships between chromosomes, genes, and alleles (Orcajo & Aznar, 2005). It has been stated that students have a limited understanding of the location, function, and structure of genes (Lewis et al., 2000a; 2000b; 2000c).

There are several studies which have focused on determining students' understanding of genetics concepts. The data collection methods employed in these studies are interviews (Marbach-Ad, 2001; Venville et al., 2005), written questions (Brown, 1990; Lewis & Wood-Robinson, 2000), multiple-choice tests (Okebukola, 1990), drawings (Rotbain et al., 2005), worksheets (Venville & Treagust, 1998), word association tests (Bahar, Johnstone, & Sutcliffe, 1999), and concept maps (Çakır & Crawford, 2001; Marbach-Ad, 2001). However, a two-tier test which was developed in relation to genetics concepts was not encountered in the literature.

Many concepts in genetics include imaginary and theoretical characteristics. Students' learning of the theoretical concepts in genetics depends on their knowing the hypothetical rationales behind these concepts (Baker & Lawson, 2001). Investigation into students' knowledge regarding the concepts of genetics, as well as the foundation on which they base their knowledge, is bound to have significant consequences in determining their understanding. Therefore, this research aims to develop a two-tier test which makes an attempt at determining secondary education students' understanding of the concepts of genetics, and to conduct validity and reliability studies of this test.

Method

The two-tier test aimed at determining students' understanding of fundamental genetics concepts was developed by using a ten-step method containing three main phases which was proposed by Treagust (1988). The steps taken in the study are as follows: (Lin, 2004; Odom & Barrow, 1995; Peterson et al., 1989; Tan et al., 2002; Treagust, 1988)

- A. Defining the content
 1. Identifying propositional knowledge statements
 2. Developing a concept map
 3. Relating propositional knowledge to the concept map
 4. Validating the content
- B. Obtaining information about students' misunderstandings
 5. Reviewing the relevant literature
 6. Interviewing students
 7. Developing multiple choice content items with free response answers
- C. Developing the two-tier diagnostic test
 8. Developing the two-tier multiple-choice test
 9. Designing a specification grid
 10. Continuing the refinements

Defining the content. This phase consisted of the first four steps which were concerned with defining the boundaries of concepts and topics. Initially, 25 propositional knowledge statements (containing every aspect of the relevant topics and concepts that are available in textbooks, supplementary books, and biology curriculum) were identified (See Appendix). These propositional statements consist of the knowledge required to understand the concepts of DNA, genes, and

chromosomes – the fundamental concepts of genetics – and the relationships between them; as well as the relationship of cell divisions with inheritance.

Concept maps were used in defining the content since they are capable of showing the key concepts of the topic and the relationships between these concepts. In the second step a concept map containing all of the concepts and the relationships of the topic was prepared by following the steps proposed by Novak (1990). Afterwards, propositional knowledge statements were associated with the concept map; and thus the internal consistency of the content was examined (Treagust, 1988). As the last step of this phase, three science educators, one field expert, and two biology teachers were asked for their expert opinions. In line with these opinions, some incoherent and incomplete statements were modified and the final format of the propositional knowledge statements and concept map was decided on. This format functioned as the basis for the test development.

Obtaining information on students' misunderstandings. The three steps that were available in this phase are related to revealing students' knowledge about the specified content. Firstly, relevant literature was reviewed and conceptual difficulties related to students' misunderstandings were determined. Students' misunderstandings about the concepts of genetics which were obtained from the literature review are presented in Table 1.

Table 1

Examples of Common Misunderstandings about Genetics and Inheritance

Only cells with the same structure or from the same part of the body contain the same genetic information.
Each cell carries different genetic information according to its function.
Genetic information is only found in gametes.
Genetic information is related to the X and Y chromosomes of sperm.
Gametes carry both members of the pairs of chromosomes found in the parent's body cells.
Sperm carry the genes responsible for half the features found in the offspring.
The gene and allele are different, or alleles are part of the gene.
A chromosome is a piece of DNA.
All chromosomes in a person carry the same genetic information.
Sex chromosomes only exist in gametes.

The data obtained from the literature review were used both in test development and in forming the interview questions. Semi-structured interviews were conducted with 21 students in order to gain a deeper perspective of students' understanding. In preparing the interview form, the current literature (Baker & Lawson, 2001; Banet & Ayuso, 2000; Johnson & Stewart, 2002; Lewis et al., 2000b, 2000c; Lewis & Wood-Robinson, 2000; Venville & Treagust, 1998; Venville et al., 2005; Wood-Robinson, Lewis, & Leach, 2000) was used. The translation and adaptation of the questions which were considered appropriate for the research were made by biology education experts, and items connected with the fundamental concepts of genetics were added to the interview form. In consequence a 10-item pre-interview form was created. To ascertain the validity of the form, expert opinion was obtained. Following the application of the pre-interview process for 2 secondary education students, one more item was added to the form. Thus, the final format for the interviews was

decided, and was made ready for application. All the interviews with students were conducted by the researcher. Consequently, students were found to partially understand, misunderstand, or to fail to understand the concepts of genetics.

A sample of the interview questions which were effective in determining students' misunderstandings follows:

- How are genetic characteristics inherited by offspring?
- What is the relationship of cell divisions with inheritance?
- What were the difficult points in learning genetics topics?

In order to obtain information about students' misunderstandings, a 13-item multiple choice test with free response answers was developed. This test was based on propositional knowledge statements, and designed according to findings gained from interviews and the literature. In this test the first tier of each item contained 5-option multiple choice questions; whereas the second tier consisted of open-ended questions which asked students the reason for the response they gave in the first tier. Prior to the application of this test to the students, expert opinion concerning the questions was consulted and the necessary adjustments were carried out. This multiple choice test with free response answers was applied to 92 students. Thus, in the second phase of the test development process, information regarding students' misunderstandings of the fundamental concepts of genetics was obtained through interviews and free response answers.

Developing the two-tier diagnostic test. In the third phase, multiple-choice two-tier test items aimed at determining students' understanding of the concepts of genetics were developed. Each item on the test was made up of two questions, where the first question was designed to measure students' knowledge of the concepts of genetics; and the second question searched for the rationale behind the option chosen by the student in the first question. The first tier of each item on the test contained a 3-option content question, whereas the second tier contained a set of five possible reasons for the answer given to the first tier question. These reasons were obtained from the findings of the interviews, the literature, and the multiple-choice test with free response answers given in the previous phase.

The data obtained from the application of the multiple-choice test with free response answers were subject to item analysis using the ITEMAN programme. Following the item analysis and the evaluation of the students' answers, two items were removed from the test, one item was re-arranged, and three items were added. Hence a two-tier test comprised of 14 items was developed.

The test content was designed in relation to target concepts to be instilled into the students. The target concepts were composed of topics which included the properties of genes, DNA, and chromosomes; their location in the organism; and their interrelations. In addition, being able to understand the relationships between cell divisions and inheritance, and being able to solve genetics problems; were other target behaviours. The items measuring the target behaviours concerning the properties of genes in the test were items 2, 6, and 8; whereas the items measuring the target behaviours concerning the concepts of chromosomes were items 3, 9, and 14. Questions 1, 4, 5, and 11 were aimed at instilling knowledge of the relationships between genes, DNA, and chromosomes. Also, the items measuring the target

behaviours concerning the relationship between cell divisions and inheritance were the 7th, 10th, 12th, and 13th items. A sample of the items used in the two-tier diagnostic test of the concepts of genetics is given in Figure 1 below:

<p>Where is the gene that determines eye colour? a) in all the cells b) in the cornea c) in the sperm</p> <p><u>Which is the reason for your answer?</u></p> <p>(1) Each gene is located in the relevant tissue. (2) The X and Y chromosomes in the sperm carry all the genes. (3) A gene determining a feature is available in all cells. (4) Each location in the body has genes specific to it. (5) The cornea is the region which is responsible for eye colour.</p> <p>Which is <u>incorrect</u> for chromosomes? a) Sex chromosomes are available only in sex cells. b) There are many genes in a chromosome. c) There is DNA in the chromosome structure.</p> <p><u>Which is the reason for your answer?</u></p> <p>(1) Sex chromosomes are available in all cells. (2) A gene is formed by the joining of chromosomes. (3) The chromosomes make up the DNA. (4) Chromosomes are in the gene because the gene has a bigger structure. (5) Sex chromosomes are available in somatic cells.</p>
--

Figure 1. A sample of items used in the two-tier genetics concepts test

A specification grid showing which propositional knowledge statements and topics each item contained was prepared. The specification grid was designed to make sure that the test was fair in that it contained the relevant propositional knowledge statements and concepts (Treagust, 1988). By conducting the test reliability and validity analyses refinements to the test items continued, so as to make the test ready to use.

Research Group

The research was conducted with 231 secondary education students. All of the students studied the topics of The Basic Components of Living Things, and The Basic Units of Living Things - Cell, Molecules Carrying Genetic Information, and Inheritance; which were in the secondary school biology curriculum and which consisted of the fundamental concepts of genetics. The age range of the research group students was between 16 and 19. Of the group, 118 (51.1%) were girls and 113 (48.9%) were boys. The research group students attended Anatolian high schools, State high schools, and high schools with an intensive foreign language programme, where the same biology curriculum had been applied.

The Analysis of the Data

The data obtained from the two-tier test application were analyzed using the Microsoft Office Excel programme and ITEMAN for Windows 3.50 in order to examine the psychometric characteristics of the instrument: reliability, discrimination

indices, difficulty indices, and the functionality of distracters. Answers were sought to two basic questions for the purposes of validity determination: first, whether or not the items measured the content that was defined by propositional knowledge statements; and second, whether or not the items were suitable for the ability level of the research group. Expert opinion was consulted for content validity, and was checked with propositional knowledge statements, as well as with concept maps. A specification grid was designed in order to determine the face validity (Odom & Barrow, 1995). For constructing validity, the method of group differences, one of the methods pointed out by Baykul (2000) in determining construct validity, was utilized. An item was scored as correct when both the desired content and the related reason were selected. Items of diagnostic instrument were evaluated for both correct and incorrect combinations selected by students. The analysis of incorrect response combinations provided data on students' misunderstandings of concepts related to that specific item.

Findings and Results

After conducting the item analysis of the data obtained by applying the two-tier test, which had been developed in relation to the fundamental concepts of genetics, the difficulty and discrimination indices of each item were calculated and presented in Table 2 below:

Table 2
The Difficulty and Discrimination Indices of Test Items

Item no	Difficulty indices	Discrimination indices
1	0.49	0.62
2	0.25	0.48
3	0.30	0.34
4	0.33	0.62
5	0.36	0.70
6	0.45	0.74
7	0.64	0.67
8	0.49	0.65
9	0.32	0.69
10	0.58	0.60
11	0.42	0.65
12	0.67	0.44
13	0.43	0.55
14	0.28	0.58
Mean	0.43	0.60

As can be seen in Table 2, the item discrimination indices were found to be between 0.34 and 0.74. These values demonstrated that the items functioned in a satisfactory way and, in particular, that they were able to distinguish between students who understood the concepts and who misunderstood them (Crocker & Algina, 1986). The item difficulty indices were between 0.25 and 0.67, and had a distribution with a large interval. The test mean was found to be 6.02 and the test difficulty to be 0.43. The test variance was 16.13. The test reliability was calculated

using the Kuder-Richardson 20 formula and was found to be 0.86. The functionality of distracters was analysed using ITEMAN for Windows 3.50. Consequently, it was found that students receiving high scores in the test had chosen the correct choices; while those receiving lower scores in the overall test had preferred distracters rather than the correct choices. The results of the analysis showed that the distracters functioned well and that there was no need for modifications in the choices. In order to determine construct validity, the correlation between the data obtained from students from Anatolian high schools and state high schools, plus high schools with an intensive foreign language programme, was calculated, and was found to be 0.169. The fact that this value was low was an indicator of test construct validity (Baykul, 2000).

The total scores gained by the students from the 14-item two-tier test ranged between 0 and 14. The number of students with zero correct answers to all 14 questions was 18; whereas the number of students who answered all the questions correctly was 4. Items of diagnostic instrument were evaluated for first tier response, and for the correct response to both tiers by students. Table 3 shows the percentage of correct answers to the first tier and to both tiers of each item.

Table 3
The Percentage of Students' Correct Answers to the First Tier and to Both Tiers of Test Items

Items	First tier	Both tiers
1	56.7	48.9
2	43.7	24.7
3	79.7	29.9
4	61.0	33.3
5	46.3	36.4
6	49.8	45.5
7	71.9	63.6
8	55.0	49.4
9	34.6	32.5
10	67.1	58.0
11	65.8	42.0
12	81.0	67.1
13	68.4	42.9
14	41.6	28.1

According to Table 3, the percentage of students giving correct answers to both tiers was lower than the percentage who gave a correct response to the first tier for each item. The percentages of correct answers to the first tier were within the 34.6 – 81.0 range; while the percentages of correct answers to both tiers were between 24.7 and 67.1. A comparison of the percentage of students who correctly answered the content part of the questions with the percentage of those who correctly answered both parts of the questions suggested that many students had learned facts without an adequate understanding of the properties and concepts involved (Peterson et al.,

1989). The graph showing the percentage of students' correct answers to the first tier and to both tiers for each test item is presented in Figure 2.

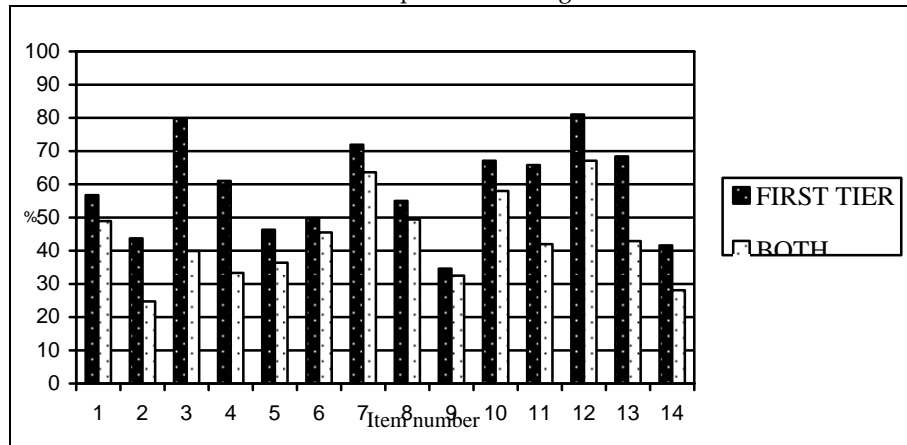


Figure 2. The percentage of students' correct answers to the first tier and to both tiers of test items

It is evident in Figure 2 that the difference between the correct answers to the first tier and both tiers, especially for items 3, 4, 11, 12, and 13, was quite high; and that the difference was low in items 6 and 9. The common feature for items 4 (61.0-33.3) and 11 (65.8-42.0) was that they were about the relationship of the concepts of DNA, genes, and chromosomes. This finding suggests that students know the concepts of DNA, genes, and chromosomes; but that they do not have a sufficient understanding of the interrelations. Similar findings were reported by Lewis et al. (2000a) who found that there was a limited awareness of the relationship between chromosomes, genes, and DNA.

Another item that deserved attention when examining the percentage differences was item 3 (79.7-29.9). This item is about the chromosome numbers, and it is an item for measuring knowledge related to the sex chromosome and the somatic chromosome. On examining the answers given to second tier of this item, it was found that students confused the concepts of sex chromosomes and somatic chromosomes.

Item 13 (68.4-42.9) was an item with a wide difference between the percentages. This was the item to measure students' understanding of relations between cell divisions and inheritance. An examination of the answers found that students did not fully understand in which cells mitosis and meiosis occurred. We again concluded that sex cells were being confused with somatic cells. The percentage difference for item 12, another item to measure knowledge concerning the relationship between cell divisions and inheritance, was also high (81.0-67.1). The responses given to this item showed that students could not distinguish between the consequences of mitosis and meiosis. Lewis and Wood-Robinson (2000) stated that most students seemed to be unaware of the nature of the difference between mitosis and meiosis; while even fewer could correctly locate mitotic division in somatic cells

and meiotic division in germ cells. The findings from their research suggest that most students didn't understand the processes and purposes of cell division; in addition to failing to make the connection between cell division and the continuity of genetic information.

The two items with the smallest difference between the percentages of correct answers given to the first tier and to both tiers were 9 (34.6-32.5) and 6 (49.8-45.5). The purpose of these items was to measure knowledge about the fact that genes which can determine a feature are available in all cells. It was understood from these items that if students knew the correct answers, they also knew the reasons for them. Banet and Ayuso (2000) confirmed that very few students thought that all cells contain the same genetic information, and very few could provide a suitable justification.

The findings also indicate that, similar to other studies that made use of the two-tier diagnostic instrument (Peterson et al., 1989; Tan et al., 2002), students performed better in the content (first tier) item than in both tiers, when the content and the reason are both considered. These consequences resulted either from the fact that students did not know the reasons underlying their knowledge or from their misunderstandings. Following the procedure used by Peterson et al. (1989), evidence of a misunderstanding was determined by looking at students' selection of an incorrect response to the content part of an item and/or an incorrect reason choice. In a similar manner to Peterson et al. (1989) and Tan (2000), students' misunderstandings were listed if they existed in at least 10% of the student sample. A higher value was not used as it might exclude some valid student misunderstandings. Students' misunderstandings of the concepts of genetics which were identified in this study are listed in Table 4 below:

Table 4
Students' Misunderstandings of Fundamental Genetics Concepts Identified by the Two-Tier Genetics Concepts Test

Misunderstandings	Percentage of Students with Misunderstanding	Item
Genes determining genetic features are found in sex cells.	55	2
Genetic features are transferred to the offspring through sex chromosomes.	48	2-14
Chromosomes determining gender are in sperm and egg cells.	40	9
The structure of DNA is composed of chromosomes.	21	1-4-5
Zygote undergoes meiosis.	19	13
The X and Y chromosomes available in sperm carry all genes.	19	6
The gene has a bigger structure than chromosomes.	14	1
Chromosomes are located on genes.	14	1
Sex chromosomes are in sex cells.	13	5-9
Genes are different from one to another in each cell.	11	11
Each gene is found in the relevant tissue.	10	6
The chromosome number is the same in each cell of an individual.	10	7

Conclusions and Discussion

The findings from the study indicate that the two-tier genetics concepts test is able to diagnose the extent of students' understanding of the concepts of genetics. The findings also show that students performed better when only the first part of the items was considered, than when both parts of the items were considered. This indicated that students selected correct content responses through rote learning without an understanding of the underlying reasons (Peterson et al, 1989).

It is generally accepted that providing correct answers does not necessarily mean the problem has been understood. Indeed, students may use some personal problem solving procedure and answer correctly without reasoning and without understanding the concepts they are using (Banet & Ayuso, 2000; Hackling & Treagust, 1984). Two-tier tests are not only able to analyze the knowledge possessed by students, but also the understanding on which they base their knowledge. Hence, an assessment to what extent students understand the concepts is made possible. As a consequence of this research, which set out to develop a two-tier test to effectively evaluate students' understanding of the concepts of genetics, a reliable and valid test was designed and was made available for use by teachers and researchers. The fact that the test uses a multiple-choice format makes it easy for teachers to apply and to score. In addition, thanks to the availability of rationale questions, teachers will be able to evaluate students' understanding of the concepts more effectively. The characteristics of the test are shown in Table 5 below:

Table 5
The Characteristics of the Two-Tier Genetics Concepts Test

Concepts and topics evaluated	:	Gene concept DNA concept Chromosome concept Relation between Gene-DNA-Chromosome Relation between cell divisions and inheritance
Number of items	:	14
Test format	:	Two-tier multiple choice: First tier - Content knowledge Second tier - Reason for student responses
Recommended grade level	:	Graduated students from high school
Time to complete test	:	25 to 35 minutes
Discrimination indices	:	Mean 0.60
		Range (items) 0.30 - 0.39 (1 item) 0.40 - 0.49 (2 items) 0.50 - 0.59 (2 items) 0.60 - 0.69 (7 items) 0.70 - 0.79 (2 items)
Difficulty indices	:	Mean 0.43
		Range (items) 0.20 - 0.29 (2 items) 0.30 - 0.39 (4 items) 0.40 - 0.49 (5 items) 0.50 - 0.59 (1 item) 0.60 - 0.69 (2 items)
Validity	:	Propositional knowledge statements, concept map, specification grid, expert opinion.
Reliability	:	0.86

Recommendations

By using two-tier diagnostic instruments at the beginning or upon completion of a specified topic, science instructors can achieve a deeper understanding of the nature of students' knowledge and the existence of any alternative conceptions or misunderstandings in the particular topic being studied (Treagust, 2006). Students' misunderstandings, which were uncovered in this study, indicate that a two-tier test may serve as an effective diagnostic tool.

It is well known that an incorrect or incomplete understanding of fundamental concepts has a negative impact on the ability to learn the related concepts and topics. Determining which concepts and topics are misunderstood by students is a vital step towards removing the obstacles which bar the way to meaningful learning. Therefore, it is believed that the use of two-tier tests, by which students' understanding of concepts is effectively evaluated, might be beneficial.

Besides making a more effective evaluation possible, two-tier tests provide ease of application and scoring. Therefore, we believe that developing and using such two-tier diagnostic tests in various fields will make significant contributions to science instruction. As mentioned by Treagust (2006), when used effectively, these tests can contribute to a deeper understanding by students of the scientific concepts covered in the curriculum.

References

- Bahar, M., Johnstone, A. H., & Hansell, M. H. (1999). Revisiting learning difficulties in biology. *Journal of Biological Education*, 33 (2), 84-89.
- Bahar, M., Johnstone, A. H., & Sutcliffe, R. G. (1999). Investigation of students' cognitive structure in elementary genetics through word association tests. *Journal of Biological Education*, 33 (3), 134-141.
- Baker, W. P. & Lawson, A. E. (2001). Complex instructional analogies and theoretical concept acquisition in college genetics. *Science Education*, 85, 665-683.
- Banet, E. & Ayuso, E. (2000). Teaching genetics at secondary school: A strategy for teaching about the location of inheritance information. *Science Education*, 84, 313-351.
- Baykul, Y. (2000). *Eğitim ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. [Measurement in Education and Psychology]. Ankara: ÖSYM Yayınları.
- Brown, C. R. (1990). Some misconceptions in meiosis shown by students responding to an advanced level practical examination question in biology. *Journal of Biological Education*, 24 (3), 182-186.
- Chen, C. C., Lin, H. S., & Lin, M. L. (2002). Developing a two-tier diagnostic instrument to assess high school students' understanding - the formation of images by a plane mirror. *Proceedings of National Science Council ROC(D)*, 12 (3), 106-121.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. USA: Holt, Reinhart and Winston.

- Çakır, M. & Crawford, B. (2001, January). *Prospective biology teachers' understanding of genetics concepts*. Annual Meeting of the Association for the Education Teachers in Science, January 18-21, Costa Mesa.
- Hackling, M. & Treagust, D. (1984). Research data necessary for meaningful review of grade ten high school genetics curricula. *Journal of Research in Science Teaching*, 21, 197-209.
- Haslam, F. & Treagust, D. F. (1987). Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument. *Journal of Biological Education*, 21 (3), 203-211.
- Johnson, S. K. & Stewart, J. (2002). Revising and assessing explanatory models in high school genetics class: A comparison of unsuccessful and successful performance. *Science Education*, 86, 463-480.
- Knippels, M. P. J., Waarlo, A. J., & Boersma, K. T. (2005). Design criteria for learning and teaching genetics. *Journal of Biological Education*, 39 (3), 108-112.
- Lewis, J., Leach, J., & Wood-Robinson, C. (2000a). All in the genes? - Young people's understanding of the nature of genes. *Journal of Biological Education*, 34 (2), 74-79.
- Lewis, J., Leach, J., & Wood-Robinson, C. (2000b). What's in a cell? - Young people's understanding of the genetic relationship between cells, within an individual. *Journal of Biological Education*, 34 (3), 129-132.
- Lewis, J., Leach, J., & Wood-Robinson, C. (2000c). Chromosomes: the missing link - young people's understanding of mitosis, meiosis, and fertilization. *Journal of Biological Education*, 34 (4), 189-199.
- Lewis, J. & Wood-Robinson, C. (2000). Genes, chromosomes, cell division and inheritance - do students see any relationship? *International Journal of Science Education*, 22 (2), 177-195.
- Lin, S. W. (2004). Development and application of a two-tier diagnostic test for high school students' understanding of flowering plant growth and development. *International Journal of Science and Mathematics Education*, 2, 175-199.
- Marbach-Ad, G. (2001). Attempting to break the code in student comprehension of genetic concepts. *Journal of Biological Education*, 35 (4), 183-189.
- Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, 27 (10), 937-949.
- Odom, A. L. & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching*, 32 (1), 45-61.
- Okebukola, P. A. (1990). Attaining meaningful learning of concepts in genetics and ecology: An examination of the potency of the concept-mapping technique. *Journal of Research in Science Teaching*, 27 (5), 493-504.
- Orcajo, T. I. & Aznar, M. M. (2005). Solving problems in genetics II: Conceptual restructuring. *International Journal of Science Education*, 27 (12), 1495-1519.
- Palmer, D. H. (1998). Measuring contextual error in the diagnosis of alternative conceptions in science. *Issues in Educational Research*, 8 (1), 65-76.

- Peterson, R. F., Treagust, D. F., & Garnett, P. (1989). Development and application of a diagnostic instrument to evaluate grade-11 and -12 students' concepts of covalent bonding and structure following a course of instruction. *Journal of Research in Science Teaching*, 26 (4), 301-314.
- Rotbain, Y., Marbach-Ad, G., & Stavy, R. (2005). Understanding molecular genetics through a drawing based activity. *Journal of Biological Education*, 39 (4), 174-178.
- Tamir, P. (1989). Some issues related to the use of justifications to multiple-choice answers. *Journal of Biological Education*, 23 (4), 285-292.
- Tan, K. C. D. (2000). *Development and application of a diagnostic instrument to evaluate students' conceptions of qualitative analysis*. PhD Thesis, Curtin University of Technology.
- Tan, K. C. D., Goh, N. K., Chia, L. S., & Treagust, D. F. (2002). Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *Journal of Research in Science Teaching*, 39 (4), 283-301.
- Tekkaya, C., Özkan, Ö., Sungur, S., & Uzuntiryaki, E. (2001). Biology concepts perceived as difficult by Turkish high school students. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 21, 145-150.
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10 (2), 159-169.
- Treagust, D. F. (2006). *Diagnostic assessment in science as a means to improving teaching, learning and retention*. (Invited presentation). UniServe Science Assessment Symposium Proceedings.
- Treagust, D. F. & Chandrasegaran, A. L. (2007). The Taiwan national science concept learning study in an international perspective. *International Journal of Science Education*, 29 (4), 391-403.
- Venville, G. J. & Treagust, D. F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. *Journal of Research in Science Teaching*, 35 (9), 1031-1055.
- Venville, G., Gribble, S. J., & Donovan, J. (2005). An exploration of young children's understandings of genetics concepts from ontological and epistemological perspectives. *Science Education*, 89, 614-633.
- Wang, J. R. (2004). Development and validation of a two-tier instrument to examine understanding of internal transport in plants and the human circulatory system. *International Journal of Science and Mathematics Education*, 2, 131-157.
- Wood-Robinson, C., Lewis, J., & Leach, J. (2000). Young people's understanding of the nature of genetic information in the cells of an organism. *Journal of Biological Education*, 35 (1), 29-36.

Appendix**Propositional knowledge statements**

1. Chromosomes are in the cell nucleus.
2. There is DNA in the chromosome structure.
3. Chromosomes are formed by coiling around proteins, thereby condensing of DNA during the cell division.
4. DNA molecules are made up of thousands of nucleotide pairs.
5. A DNA molecule contains many genes.
6. Gene is a certain length part of DNA.
7. Genes are located on chromosomes.
8. Genes determine all genetic features.
9. Genes are available in all cells.
10. A gene that is responsible for a feature is available in all cells, not only in the cell related with that feature.
11. Many genes are available in a chromosome.
12. The most important function of chromosomes is to carry the genes.
13. Chromosomes in a living organism are composed of sex chromosomes and somatic chromosomes.
14. The number of chromosomes in human beings is 46, 44 of which are somatic chromosomes.
15. The sex chromosomes in human beings are XY in men and XX in women.
16. Sex chromosomes determine gender.
17. Both sex chromosomes and somatic chromosomes are available in all cells.
18. The transfer of parents' features into zygote by joining of fathers' and mothers' sex cells is called inheritance.
19. The number of chromosome is haploid in sex cells, diploid in somatic cells.
20. There are same genes in all somatic cells of a living organism.
21. Sex cells are formed by meiosis.
22. There are different genes in sex cells of living organisms because of crossing over.
23. While the number of cells increases through mitosis, the number of chromosomes remains constant.
24. A multi-cell mature organism develops from a one-cell zygote through mitosis. Therefore, the cells in the organism's body have chromosomes and genes of the same number and the same type.
25. The number of chromosomes is reduced to half by meiosis.

Öğrencilerin Genetik Kavramları Anlamalarını Belirlemeye Yönelik İki Aşamalı Bir Testin Geliştirilmesi

(Özet)

Problem Durumu: Öğrencilerin fen kavramlarını eksiksiz ve doğru olarak anlamalarını amaçlayan fen öğretiminin etkililiği, ancak uygun ölçme araçları ile ölçüldüğünde doğru olarak değerlendirilebilir. Öğrencilerin fen kavramlarını anlamalarını belirlemede kullanılan kavram haritaları, çoktan seçmeli testler ve görüşmeler başarılı yöntemler olarak kabul edilmektedir ancak bu yöntemlerin her birinin çeşitli sınırlılıkları mevcuttur. Öğrencilerin, kavramları anlamalarının belirlenmesinde, öğrenci cevaplarının nedenlerine dayanan ve bilinen alternatif kavramları içeren iki aşamalı çoktan seçmeli test maddelerinin kullanılması önerilmektedir. İki aşamalı testler, öğrencilerin seçtikleri cevaplara ait gerekçelerini ve yorumlarını dikkate aldığından önceki yaklaşımlara göre oldukça gelişmiş bir yaklaşımdır. Puanlaması ve uygulaması kolay olan iki aşamalı testler ile geçerli ve güvenilir biçimde, öğrenci fikirleri daha etkili olarak değerlendirilebilmektedir.

Öğrencilerin anlamada en fazla zorlandığı biyoloji konularından biri genetikdir. Çoğu genetik kavramı imgesel ve teorik bir özelliğe sahiptir. Öğrencilerin teorik genetik kavramları anlamlı olarak öğrenebilmeleri, bu kavram ve olguların varsayıma dayanan nedenlerini bilmelerine bağlıdır.

Araştırmanın Amacı: Öğrencilerin genetik kavramları anlamalarını belirlemede, öğrencilerin bu kavramlar hakkında sahip oldukları bilgiler ile beraber bu bilgilerini dayandırdıkları nedenlerin araştırılması önemli ve etkili sonuçlar verecektir. Bu sebeple, araştırmada ortaöğretim öğrencilerinin genetik kavramları anlamalarını belirlemeye yönelik iki aşamalı bir testin geliştirilmesi, geçerlik ve güvenilirlik analizlerinin yapılması amaçlanmıştır.

Araştırmanın Yöntemi: Ortaöğretim öğrencilerinin genetiğin temel kavramlarını anlamalarını belirlemeye yönelik iki aşamalı test, 10 aşamalı 3 ana bölümden oluşan bir yöntem izlenerek geliştirilmiştir. Birinci bölüm içeriğin belirlenmesi, ikinci bölüm öğrencilerin yanlış anlamaları hakkında bilgi edinilmesi, üçüncü bölüm ise iki aşamalı testin geliştirilmesidir. İçeriğin belirlendiği ilk bölümde bilgi önermeleri ve kavram haritasıyla konunun sınırları çizilmiş, kapsam geçerliği sağlanmıştır. İkinci bölümde öğrencilerin yanlış anlamalarını belirlemek için ilgili araştırmalar incelenmiş, öğrencilerle görüşmeler yapılmış, açık uçlu ve çoktan seçmeli sorular içeren bir test uygulanmıştır. Üçüncü bölümde, iki aşamalı çoktan seçmeli test maddeleri geliştirilmiştir.

Geliştirilen iki aşamalı testin 231 ortaöğretim öğrencisine uygulanmasıyla elde edilen verilerin, madde ve test analizi yapılmıştır. Analizler sonucunda maddelerin güçlüğü, ayırt edicilikleri, çeldiricilerin işlerliği ve testin güvenilirliği belirlenmiştir. Geçerliğin belirlenmesi için iki temel soruya cevap aranmıştır. Bunlardan ilki; maddelerin bilgi önermeleri ile tanımlanan içeriği ölçüp ölçmediği, ikincisi ise; maddelerin uygulama grubunun düzeyine uygun olup olmadığıdır. Kapsam geçerliği için uzman görüşleri alınmış, bilgi önermeleri ve kavram haritası ile kontrolü sağlanmıştır. Yapı geçerliğinin saptanmasında grup farklılıkları yöntemi kullanılmıştır.

Bir sorunun doğru cevaplanmış olarak kabul edilebilmesi için, maddenin her iki aşamasına da doğru cevap verilmiş olması gerekmektedir. Öğrencilerin test maddelerinin hem birinci aşamasına hem de her iki aşamasına birden verdikleri cevaplar değerlendirilmiştir. Yanlış cevap kombinasyonları incelenmiş ve öğrencilerin genetik kavramları ile ilgili yanlış anlamaları belirlenmiştir.

Araştırmanın Bulguları: İki aşamalı genetik kavram testinin uygulanmasıyla elde edilen verilerin madde analizi sonucunda maddelerin ayırtıcılık güçleri, 0,34 ile 0,74 arasında bulunmuştur. Madde güçlük indeksleri ise 0,25 ile 0,67 arasında değişmektedir. Testin ortalaması 6,02, testin ortalama güçlüğü ise 0,43 olarak hesaplanmıştır. Testin varyansı ise 16,13'tür. Testin güvenilirliği Kuder-Richardson 20 formülü ile hesaplanmış ve 0,86 olarak bulunmuştur. ITEMAN for Windows 3,50 ile çeldiricilerin işlerliği analiz edilmiştir. Analiz sonucunda testten yüksek puan alan öğrencilerin, maddelere ait doğru seçenekleri cevaplarırken, testin genelinden düşük başarı elde edenlerin ise doğru seçenek dışındaki çeldiricilere yöneldikleri görülmüştür. Analiz sonuçları her bir çeldiricinin iyi işlediğini, seçeneklerde düzeltmeye gerek olmadığını göstermiştir. Yapı geçerliğini belirlemek için grup farklılıkları yöntemi kullanılmıştır. Anadolu Lisesi öğrencileri ve düz lise ile yabancı dil ağırlıklı lise öğrencileri iki farklı grup olarak değerlendirilmiş, bu gruplardan elde edilen veriler arasındaki korelasyona bakılmıştır. Korelasyon katsayısı 0,169 olarak bulunmuştur. Bu değerin düşük olması testin yapı geçerliğinin bir göstergesi olarak kabul edilmektedir.

Öğrencilerin test maddelerinin birinci aşamasına doğru cevap verme oranlarının, her iki aşamaya birden doğru cevap verme oranlarından yüksek olduğu belirlenmiştir. Bu bulgu, öğrencilerin kavramları yeterince anlamadan, ezberleyerek öğrendiklerini göstermektedir. Yanlış cevaplar incelenmiş ve öğrencilerin %10'undan fazlasında karşılaşılan yanlış anlamalar belirlenmiştir.

Araştırmanın Sonuçları ve Önerileri: İki aşamalı genetik kavram testinin madde analizi, geçerlik ve güvenilirlik analizleri sonucunda elde edilen tüm bulgular, bu testin ortaöğretim öğrencilerinin genetik kavramları anlamalarını belirlemek amacıyla geçerli ve güvenilir biçimde kullanılabilceğini göstermektedir. Geliştirilen testte yer alan maddelerin ilk aşaması 3 seçeneikli bir bilgi sorusundan oluşurken, ikinci aşaması, birinci aşamada verilen cevap için 5 muhtemel nedenden oluşan seçenekler içerir. Test, genetiğin temel kavramları olan DNA, gen, kromozom kavramları ile bu kavramlar arasındaki ilişkileri ve hücre bölünmelerinin kalıtımla ilişkisi konularında öğrencilerin anlamalarını belirlemeye yöneliktir. Test, iki aşamalı 14 çoktan seçmeli maddeden oluşmaktadır. Testin çoktan seçmeli oluşu sınıf içerisinde öğretmenlere uygulama ve puanlama kolaylığı sağlarken, öğrencilere cevaplarının nedenlerinin sorulması ile öğrencilerin kavramları anlamaları daha etkili olarak değerlendirilebilecektir.

Anahtar Sözcükler: İki aşamalı test, genetik kavramlar, yanlış anlamalar.

Views of Primary School Administrators, Teachers and Parents on Parent Involvement in Turkey

Ali Sabancı*

Suggested Citation:

Sabancı, A. (2009). Views of primary school administrators, teachers and parents on parent involvement in Turkey. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, 36, 245-262.

Abstract

Problem Statement: The way teachers and parents of school-aged children interact has a major impact on establishing an effective educational environment and on learning. Parents are one of the most important variables that are expected to contribute to schools voluntarily both in class activities and in managerial processes. Parents have traditionally adopted the role of supporter or representative, rather than becoming full and equal educational partners.

Purpose of Study: The purposes of this study were first to determine the administrators', teachers' and parents' attitudes towards building more effective partnerships between schools and parents in Turkey, and secondly to find out how their attitudes were affected by gender, work experience, academic background, *teaching branch*, parents' profession, parents' income and class grade of students.

Methods: This research was conducted as a survey using a descriptive method. The data was gathered using a questionnaire developed by the researcher. The reliability of the questionnaire applied to teachers and managers was found to be $\alpha=.95$. The reliability of the questionnaire applied to parents was $\alpha=.91$. Specific descriptive analyses that were conducted to calculate the data included frequencies, means, significance tests, MANOVA and one-way ANOVA.

Findings and Results: Of the groups surveyed, managers hold more positive attitudes towards parental involvement. Teachers' and parents' scores are lower in all dimensions compared to managers. Examination of the coefficients for the linear combinations distinguishing teachers' views indicated that communicating, volunteering, learning at home and collaborating contributed most to distinguishing teachers' attitudes. No significant difference was found regarding demographic variables between principals and assistant principals. There were also no differences between parents' attitudes in relation to demographic variables. In particular, communicating, volunteering, learning at home, and collaborating contributed significantly toward discriminating associate degree graduate teachers from the other two groups (bachelor's and master's degrees).

* Ass. Prof. Dr., Akdeniz University, Faculty of Education, alisabanci@akdeniz.edu.tr

Conclusions and Recommendations: The educational implication of the findings is that effective parent-school cooperation requires a revolutionary outlook with a focus on finding creative strategies and methods to transform the reported positive attitudes into practice.

Keywords: Parent involvement, parent-school partnership, student achievement, effective school

As a result of technological, social and economical transformation in the “information age,” schools cannot be seen as isolated parts where the focus is on teachers. One of the most important parts of the new perspective on education is to involve parents in decision-making and get their volunteered support into our classrooms. The way teachers and parents of school-aged children interact has a major impact on establishing an effective educational environment and learning. Parents are one of the most important variables that are expected to contribute to school voluntarily both in class activities and in managerial processes. Good relations can be thought to be a starting point but, unfortunately, studies in Turkey show that teacher-parent meetings are not effective enough to set up such good relations (Demirbulak, 2000). Lack of a reasonable relationship causes teachers to think that parents are the weakest link in the education chain, and that they always have unreasonable demands. On the other hand, in their relations with teachers, parents always experience frustration (Lasky, 2000). Indeed, a negative and unstructured relationship between a school and parents is likely to negatively influence educational operations.

Parents have traditionally adopted the role of supporter or representative, rather than becoming full and equal educational partners. Many schools still consider parents as intruders into their territory and are not fully receptive to parental suggestions and participation (Henry, 1996). Cibulka (1994) and Lasky (2000) summarize the traditional view of parental involvement in the following items: (1) teachers and other school personnel are experts who know more about how to educate children than parents; (2) educators embody a broader and more balanced view of a child's educational needs than do parents; (3) educators should be empowered to fulfil this responsibility with wide discretion and autonomy, based on professional expertise and judgement; and (4) the school may delegate tasks to the parents, such as assisting in homework and serving as volunteers for the classroom.

In recent years, conventional thinking has moved towards the view that schools cannot do their work in isolation from parents, community agencies, industry, business and universities (Henry, 1996). It is clear that not only are schools important to parents and families, but that schools need the support of parents to achieve optimum success (Berger, 1991; Dempsey & Sandler, 2000). Epstein (1995) emphasizes that the way schools care about children is reflected in the way schools care about the children's families.

In most cases there are barriers that may deter families from participating in their children's schooling, which arguably have universal relevance. These barriers include the following: (1) involvement with families and communities is initiated by

individual teachers rather than being addressed by the whole school (Epstein, 2001; Jordon, Orozco & Averett, 2001; Voorhis & Sheldon, 2004); (2) schools contact parents only when they want something; (3) insufficient time is allowed at parents' meetings for parents to discuss their child's progress with the teacher; (4) it is difficult to make private appointments to see the head or teacher; (5) there is low awareness and visibility of the school's governing body (Epstein & Dauber, 1991; Dean, 2001; Miedel & Reynolds, 1999); and (6) notes and memos are transmitted through children, rather than by direct communication (Becker & Epstein, 1982; Greenwood & Hickman, 1991).

Research concerning parental involvement suggests that there is a strong positive relationship between parental involvement and student achievement (Bagin, D., Gallagher, D. R. & Kindred, 1994; Dearing, McCartney, Weiss, Kreider, & Simpkins, 2004; Epstein, 1984; Greenwood & Hickman, 1991; Martinez, Martinez & Perez, 2004; Ng, 1999; Werf, Creemers, & Guldmond, 2001; Zellman & Waterman, 1998). Initially, the socioeconomic status of the family was considered an unchangeable factor concerning student performance, but in the 1960s and early 1970s, a common view arose that parents could contribute to educational success in schools by providing a stimulating home environment (Epstein, 1987; Miedel & Reynolds, 1999; Werf *et al.*, 2001).

Administrative Arrangements and Literature Review in Turkey

In Turkey, there is a significant struggle to facilitate parental involvement through (1) school-family associations (which can be defined as a school board, the members of which are managers, teachers and parents) that are responsible for school development (Resmi Gazete, 1961; Resmi Gazete, 2005a); (2) class-family associations that represent school-family associations at the class level (all parents are members of this association) (Resmi Gazete, 2003); (3) social activity groups in which parents participate voluntarily (Resmi Gazete, 2005b); and (4) written home-school contracts, developed in discussion with parents, whereby school, parental and student assignments are clarified (M.E.B, 2005). However, not many parents are interested in these arrangements. In fact, there are parental complaints about the lack of structured and systematically-followed involvement programmes. Parental involvement can be said to be restricted to teacher-parent meetings held once each semester, during which failure and problematic situations are stressed by the teacher; parent-initiated requests to be informed of student's success and school attendance; individual contact made by teachers in which problematic situations and in-class support of pupils tend to be the main issues; and finally, financial needs announced by managers.

Bilgin (1991) found that legal arrangements regarding parent-school relationships were adopted by all parties. Administrators and parents reported that the duties assigned were successfully applied and that parents were efficient enough on their part. However, teachers had negative views regarding the degree to which parents were efficient and carried out their duties. Günkan (2007) reported that most of the participants in her study defined parental involvement as intentional cooperation

between schools and families towards a common goal. Çelik (2005) found the most common problems reported by teachers to be a) parental concern focused exclusively on their child's academic achievements to the exclusion of their social and emotional attitudes, b) school visits by parents only when a problem occurred, and c) disregard of negative issues relating to their children. On the other hand, families complained about a) the lack of teacher visits, b) few opportunities extended to parents to participate in making collective decisions, and c) being contacted only when there is a problem with their children. Yılmaz (2006) found that families in general do not put forth sufficient effort in responding to teacher expectations to provide help and support to their children. Yıldırım and Dönmez (2008) concluded that neither schools nor parents alone are effective in meeting the educational needs of children. One-to-one teacher interviews with parents are regarded to be relatively more effective than other methods and techniques. Perceived restrictions regarding parental involvement often occur because parents do not regularly attend school, are usually petitioned only for financial help, and sometimes perceive interactions with teachers as being distant and harsh. Çelenk (2003; 2003a) found that primary school students had high levels of success and achievement when their parents had close relations with the school and adopted supportive attitudes towards education. Kebeci (2006) found knowledge of and participation in parent-school association assemblies by parents to be comparatively higher than that of teachers. However, parents' relative level of concern was found to be lower compared to teachers'.

Types of Involvement

Williams and Chavkin (1989) suggest six parental roles: audience, home tutor, programme supporter, co-learner, advocate, and decision-maker. Cervone and O'Leary (1982) proposed four types of parental involvement: reporting progress, attending special events, becoming educated, and teaching. Greenwood and Hickman (1991) discussed five types of parental involvement: acting as audience, volunteering, teaching their own children, learning, and decision-making. Castro *et al.* (2004) proposed parental involvement in five main ways—through communicating with teachers, being teachers themselves, supporting activities, learning, and advocating. According to Hill and Taylor (2004), volunteering in the classroom, communicating with the teacher, participating in academic-related activities at home, communicating the positive value of education, and participating in the parent-teacher relationship are all crucial to parental involvement. As Berger (1991) points out, parental participation in schools needs to be at five levels of involvement, from an active partner to a passive supporter: through being an active partner and educational leader at home, as a decision-maker, a school volunteer, a liaison between home and school to support homework, and as a supporter of the educational goals of the school. In addition, Epstein (1995), Voorhis and Sheldon (2004) state a framework of six major types of involvement which has evolved from many studies and from many years of work by educators and families in elementary, middle, and high schools: parenting, communicating, volunteering, learning at home, decision-making, and collaborating with the community. Parenting is helping all families establish supportive home environments for children. Communicating involves establishing two-way exchanges about school programmes and children's

progress. Volunteering encompasses recruiting and organizing parental help at school, home, or other locations. Decision-making needs to include parents from all racial, ethnic, socio-economic, and other groups in school decisions and in developing parent leaders and representatives. Training needs to be offered to enable these leaders to serve as representatives of other families. The development of parenting skills can help all families establish supportive home environments for children. Collaboration can be encouraged through identifying and integrating community-based services and resources to strengthen school programmes, family practices, and student learning and development. Learning at home can involve schools providing information to families about how to help students with homework and other curriculum-related materials.

Aims

The purposes of this study was first to determine administrator, teacher and parental attitudes towards building more effective partnerships between schools and parents in Turkey, and secondly to find out how these attitudes were affected by gender, work experience, academic background, teaching branch parents' profession, parents' income and class grade of students. The findings will help us understand underlying factors so that we can rearrange pre-service and in-service training for both managers and teachers to change their attitudes. Moreover, determining unfavourable factors about families will help us understand how to help them change their attitudes and overcome unfavourable factors to involve them more effectively.

The specific questions addressed for this purpose are:

1. What are administrator, teacher and parental attitudes towards building more effective partnerships between schools and parents? Are there differences between the attitudes of these groups regarding a linear combination of communication, volunteering, learning at home, parenting, collaborating and decision making?
2. Are there differences in administrators' attitudes towards building partnerships between schools and parents in relation to a) gender, b) work experience, c) academic background and d) *teaching branch* variables on a linear combination of communication, volunteering, learning at home, parenting, collaborating and decision making?
3. Are there differences in teachers' attitudes towards building partnerships with parents in relation to a) gender, b) work experience, c) academic background and d) *teaching branch* variables on a linear combination of communication, volunteering, learning at home, parenting, collaborating and decision making?
4. Are there differences in parents' attitudes towards building partnerships between schools and parents in relation to a) gender, b) parents profession, c) class grade of the student, d) parents' income and d) academic background?

Method

Research Design

This research was conducted as a survey using a descriptive method to ascertain administrator, teacher and parental attitudes towards building more effective partnerships between schools and parents and to draw wider implications for effective parent-school partnerships.

Population and Sample

The research population consisted of 126 state primary school principals, 199 assistant principals and 4425 teachers working in primary schools in Antalya in Turkey. Parents of the students attending the chosen primary schools comprised the parent population. In order to adequately cover the geographically dispersed educational population, multi-stage cluster sampling was used (Anderson, 1990). Initially, the total population of the research in Antalya province was divided into seven geographical educational regions. In the second stage, a random sampling technique was applied to the elements from each of the clusters; 35 primary schools as representatives of the seven geographical education regions were selected. 103 administrators of those 35 schools were selected as the sample. In order to adequately cover a population of 1010 teachers in these 35 schools, 277 primary school teachers were chosen by a random sampling technique. Parents of three students attending the 1st, 2nd, 3rd, 4th, 5th, 6th, 7th and 8th grade classes of the 35 schools were selected randomly from each grade, and 840 parents were sent questionnaires (Hair, Anderson, Tatham, & Black, 1998; Muijs, 2004). Of these, 85 (82.5%) of the administrators from a total of 103, 273 (98.5%) of the teachers from a total of 277, and 608 (72.3%) of the parents from a total of 840 completed questionnaires.

Instruments and Data Analyses

Data was gathered using a questionnaire developed by the researcher. The questionnaire had two parts. Firstly, the questionnaire gathered personal information about the participants. Secondly, it was designed to examine attitudes towards building better partnerships between parents and Turkish schools. Responses were based on a Five-Point Likert Scale ranging from "I definitely don't agree" to "I definitely agree." The questionnaire was reviewed by academicians in the field of educational administration, practicing principals, assistant principals, teachers and parents to ascertain the clarity and validity of each question, and was revised according their feedback. Loadings of less than .45 were omitted to improve clarity. After rotation, 59.960% of the variance was accounted for by six factors with 35 items for principals, assistant principals and teachers; and 60.727% of the variance was accounted for by six factors with 23 items for parents. These items were divided into the six broad thematic sections of communicating, volunteering, learning at home, parenting, collaborating and decision making.

The reliability of the dimensions (Cronbach Alpha) for the parental involvement questionnaire as applied to principals, assistant principals and

teachers was as follows: communicating: $\alpha=.90$; volunteering: $\alpha=.91$; learning at home: $\alpha=.79$; parenting: $\alpha=.70$; collaborating: $\alpha=.63$; and decision making: $\alpha=.65$. The reliability of the questionnaire as a whole was found to be $\alpha=.95$. The reliability of the dimensions of the questionnaire as applied to parents were as follows: communicating: $\alpha=.67$; volunteering: $\alpha=.70$; learning at home: $\alpha=.69$; parenting: $\alpha=.68$; collaborating: $\alpha=.76$; and decision making: $\alpha=.58$. The reliability of the questionnaire as a whole was $r=.91$. The SPSS computer program was used to analyze the relationships between the variables. Specific descriptive analyses that were conducted to calculate data included frequencies, means, significance tests, MANOVA and one-way ANOVA (Hair *et al.*, 1998).

Findings and Results

Demographic data are presented first, in Table 1, followed by results from the research questions.

Table 1

Demographic Data of Administrators and Teachers

		Administrators (N=85)		Teachers (N=273)	
		N	%	N	%
Gender	Female	19	22.4	161	59.0
	Male	66	77.6	112	41.0
Work experience	1-5 years	13	15.3	31	11.4
	6-10 years	16	18.8	66	24.2
	11-15 years	15	17.6	49	17.9
	16-20 years	19	22.4	56	20.5
	21 years +	22	25.9	71	26.0
Academic background	Associate degree	26	30.6	71	26.0
	Bachelor's degree	59	69.4	190	69.6
	Master's degree	-	-	12	4.4
Teaching branch	Class teacher	57	67.1	148	54.2
	Subject teacher	28	32.9	125	45.8

As seen in Table 1, most of the managers were male 66 (77.6%), while most of the teachers were female 161 (59%). The respondents were distributed across all levels of work experience. There were no responses to the academic background variable from managers, so no comparison could be made for this category. The distribution of respondents to the teaching branch variable in both the manager and teacher categories also provided quite usable data for analysis.

Table 2

Demographic Data of Parents

	Parent (N=608)			Parent (N=608)			
	N	%		N	%		
Profession	Unofficial worker	119	19.6	Class of student	Grade 1	89	14.6
	Self-employed	74	12.2		Grade 2	82	13.5
	Official worker/officer	100	16.4		Grade 3	77	12.7
	Official manager	10	1.6		Grade 4	86	14.1
	Housewife	243	40.0		Grade 5	82	13.5
	Retired	29	4.8		Grade 6	57	9.4
	Farmer	32	5.3		Grade 7	74	12.2
	Dead/divorced	119	19.6		Grade 8	61	10.0
	Income	0-500	185		30.4	Academic background	Elementary school
501-1000		224	36.8	Secondary school	86		14.1
1001-1500		91	15.0	High school	196		32.2
1501-2000		65	10.7	Associate degree	61		10.0
2001+		43	7.1	Bachelor's degree	69		11.3
Gender	Female	344	56.6	Master's degree	7		1.2
	Male	264	43.4				

As seen in Table 2, 344 (56.6%) of the participants were female while only 243 (40%) were housewives. Another significant demographic finding was the number of dead/divorced parents (119, or 19.6%). Of the participants, 67.2% were in a low income group. The demographic findings show that most of the parents had primary and secondary school education.

1. *Administrator, teacher and parental attitudes towards building more effective partnerships between schools and parents*

According to Table 3, managers hold more positive attitudes towards building partnerships with parents in communicating ($\bar{x}=4.1465$), volunteering ($\bar{x}=4.0225$) and decision making ($\bar{x}=4.0392$), but they hold less positive attitudes towards building partnerships when compared with parents in learning at home ($\bar{x}=3.7647$), parenting ($\bar{x}=3.9961$) and collaborating ($\bar{x}=3.8118$). Teachers' and parents' scores in communicating are lower in all dimensions compared to managers ($\bar{x}=3.9833$; 3,5201), volunteering ($\bar{x}=3.8352$; 3,8076), learning at home ($\bar{x}=3.3868$; 3,6336), parenting ($\bar{x}=3.6654$; 3,7862), collaborating ($\bar{x}=3.4121$; 3,7355) and decision making ($\bar{x}=3.8730$; 3,7311). The results show that, although teachers and parents reported positive attitudes towards building partnerships, there might be slight differences in their approach to parental involvement.

A multivariate analysis of variance was conducted to assess whether there were differences between managerial, teacher and parental attitudes on a linear combination of communication, volunteering, learning at home, parenting,

Table 3

Attitudes of Respondents towards Building Partnerships with Parents

Involvement Dimension	Position	N	\bar{X}	S	df	F	p
Communicating	Administrators	85	4.1465	.5471	1-356	4.46	.035
	Teachers	273	3.9833	.6429			
	Parents	608	3.5201	.7355			
Volunteering	Administrators	85	4.0225	.5387	1-356	5.64	.018
	Teachers	273	3.8352	.6612			
	Parents	608	3.8076	.7704			
Learning at Home	Administrators	85	3.7647	.7679	1-356	16.43	.000
	Teachers	273	3.3868	.7449			
	Parents	608	3.6336	.7832			
Parenting	Administrators	85	3.9961	.6637	1-356	13.01	.000
	Teachers	273	3.6654	.7592			
	Parents	608	3.7862	.7458			
Collaborating	Administrators	85	3.8118	.7479	1-356	13.63	.000
	Teachers	273	3.4121	.9062			
	Parents	608	3.7355	.7641			
Decision Making	Administrators	85	4.0392	.5919	1-356	3.72	.054
	Teachers	273	3.8730	.7215			
	Parents	608	3.7311	.7826			

collaborating, and decision making. A significant difference was found: Wilk's lambda = .728, $F(957, 1914) = 27.42$, $p < .000$, Multivariate $\eta^2 = .147$.

An examination of the coefficients for the linear combinations distinguishing managers', teachers' and parents' views indicated that communicating (.627), volunteering (.215), learning at home (-.246), parenting (.210), collaborating (-.323) and decision making (.309), contributed significantly toward discriminating between managers, teachers and parents ($p = 0.000$, $p = .010$, $p = 0.000$, $p = 0.015$, $p = 0.000$ and $p = 0.000$).

Follow-up univariate ANOVAs indicated that communicating, volunteering, learning at home, parenting, collaborating and decision making were significantly different for managers, teachers and parents: $F(2-962) = 60.48$, $p < .000$; $F(2-962) = 3.29$, $p < .037$, $F(2-962) = 12.40$, $p < .000$, $F(2-962) = 6.77$, $p < .001$, $F(2-962) = 16.97$, $p < .000$ and $F(2-962) = 8.18$, $p < .000$, respectively.

2. Differences between administrators' attitudes towards building partnerships between schools and parents in relation to a) gender, b) work experience, c) academic background and d) teaching branch variables on a linear combination of communication, volunteering, learning at home, parenting, collaborating and decision making

A multivariate analysis of variance was conducted to assess whether there were differences between administrators' attitudes in relation to gender, work experience, academic background and teaching branch variables on a linear combination of

communication, volunteering, learning at home, parenting, collaborating and decision making. No significant difference was found in relation to gender: Wilk's lambda= .958, $F(78, 78) = .574$, $p = .75$, Multivariate $\eta^2 = .042$. No significant difference was found in relation to work experience: Wilk's lambda= .634, $F(75, 262) = 1.526$, $p = .059$, Multivariate $\eta^2 = .108$. No significant difference was found in relation to academic background: Wilk's lambda= .910, $F(78, 78) = 1.289$, $p = .272$, Multivariate $\eta^2 = .090$. No significant difference was found in relation to teaching branch: Wilk's lambda= .946, $F(78, 78) = 1.289$, $p = .621$, Multivariate $\eta^2 = .054$.

3. *Differences between teachers' attitudes towards building partnerships with parents in relation to a) gender, b) work experience, c) academic background and d) teaching branch variables on a linear combination of communication, volunteering, learning at home, parenting, collaborating and decision making*

A multivariate analysis of variance was conducted to assess if there were differences between teachers' attitudes in relation to gender, work experience, academic background and teaching branch variables on a linear combination of communication, volunteering, learning at home, parenting, collaborating and decision making. No significant difference was found in relation to gender: Wilk's lambda= .978, $F(266, 266) = .993$, $p = .430$, Multivariate $\eta^2 = .022$. No significant difference was found in relation to work experience: Wilk's lambda= .880, $F(263, 918) = 1.433$, $p = .082$, Multivariate $\eta^2 = .032$. No significant difference was found in relation to *teaching branch*: Wilk's lambda= .964, $F(266, 266) = 1.659$, $p = .131$, Multivariate $\eta^2 = .036$. however, a significant difference was found in relation to academic background. The analysis is given in table 4.

A multivariate analysis of variance was conducted to assess whether there were differences between teachers' attitudes on a linear combination of communication, volunteering, learning at home, parenting, collaborating, and decision making in relation to academic background. A significant difference was found: Wilk's lambda= .924, $F(265, 530) = 1.784$, $p < .048$, Multivariate $\eta^2 = .039$.

Examination of the coefficients for the linear combinations distinguishing teachers' views indicated that communicating (.662), volunteering (.472), learning at home (.637) and collaborating (.690) contributed significantly toward discriminating associate degree graduates from the other two groups (bachelor's and master's degree graduates) ($p = .001$, $p = .021$, $p = .006$ and $p = .014$). Parenting and decision making did not contribute significantly to distinguishing the groups.

Follow up univariate ANOVAs indicated that communicating, volunteering, learning at home and collaborating were significantly different for teachers: $F(2-270) = 6.359$, $p < .002$; $F(2-270) = 4.047$, $p < .019$, $F(2-270) = 6.693$, $p < .001$, and $F(2-270) = 5.457$, $p < .005$, respectively.

4. *Differences between parents' attitudes towards building partnerships between schools and parents in relation to a) gender, b) parental profession, c) class grade of the student, d) parents' income and d) academic background*

Table 4***Attitudes of Teachers in Relation to Academic Background***

Involvement Dimension	Academic Background	N	\bar{X}	S	df	F	p
A- Communicating	Associate degree	71	4.1472	.53725	2-270	6.359	.002
	Bachelor's degree	190	3.9536	.61903			
	Master's degree	12	3.4848	1.15730			
B- Volunteering	Associate degree	71	4.0026	.62058	2-270	4.047	.019
	Bachelor's degree	190	3.7919	.63781			
	Master's degree	12	3.5303	1.02694			
C- Learning at Home	Associate degree	71	3.6366	.67976	2-270	6.693	.001
	Bachelor's degree	190	3.3179	.73000			
	Master's degree	12	3.0000	.98719			
D- Parenting	Associate degree	71	3.7793	.73451	2-270	1.187	.307
	Bachelor's degree	190	3.6316	.75511			
	Master's degree	12	3.5278	.94771			
E-Collaborating	Associate degree	71	3.6901	.87980	2-270	5.457	.005
	Bachelor's degree	190	3.3342	.86992			
	Master's degree	12	3.0000	1.26131			
F- Decision Making	Associate degree	71	4.0094	.67840	2-270	2.157	.118
	Bachelor's degree	190	3.8368	.71905			
	Master's degree	12	3.6389	.92614			

A multivariate analysis of variance was conducted to assess if there were differences between parents' attitudes in relation to gender, parental profession, class grade of the student, parents' income and academic background variables on a linear combination of communication, volunteering, learning at home, parenting, collaborating, and decision making. No significant difference was found in relation to gender: Wilk's lambda=.986, $F(601, 601) = 1.446$, $p = .195$, Multivariate $\eta^2 = .014$. No significant difference was found in relation to parents' profession: Wilk's lambda=.933, $F(596, 2619) = 1.150$, $p = .249$, Multivariate $\eta^2 = .011$. No significant difference was found in relation to class grade of the student: Wilk's lambda=.948, $F(595, 2794) = 1.433$, $p = .866$, Multivariate $\eta^2 = .009$. No significant difference was found in relation to parents' income: Wilk's lambda=.958, $F(598, 2087) = 1.078$, $p = .361$, Multivariate $\eta^2 = .011$. No significant difference was found in relation to academic background: Wilk's lambda=.938, $F(597, 2390) = 1.278$, $p = .144$, Multivariate $\eta^2 = .011$.

Conclusions and Recommendations

This article has focused firstly on determining primary administrator, teacher and parental attitudes towards building more effective partnerships with parents in schools in Turkey and secondly on how these attitudes were affected by occupational, educational, economical and physical factors such as gender, work experience, academic background, *teaching branch*, parents' profession, parents' income and class grade of students variables on a linear combination of communication, volunteering, learning at home, parenting, collaborating and decision making.

Conclusions drawn from demographic findings show that the number of female managers (22.4%) does not equal the number of female teachers (59.0%). The data reveals a problematic area of why women are not represented proportionately to their population in management positions in the education sector. Of the managers, 26 (30.6%) were chosen from among associate degreed teachers. The higher the level of education a manager had the more professional he/she is expected to be in managerial competencies. The ratio of associate degreed managers to teachers might cause problems in motivating, communicating and other management processes where most of the teachers in the school are in possession of a bachelor's or a master's degree. Teachers' scores are lower in communicating, volunteering, learning at home, parenting, collaborating and decision-making when compared to those of managers. The results showed that though teachers reported positive attitudes towards building partnerships with parents, there might be slight differences in their approach and how they seek to meet the needs of participating parents. An examination of the coefficients showed that communicating, volunteering, learning at home, parenting and collaborating contributed most to distinguishing the groups. One important reason for this might be lack of practice, as Epstein and Dauber (1991) ascertain, or methodological problems—that is, a lack of knowledge about how to help parents be mentors in educational matters, which mostly requires professional training. On the other hand, both groups agreed that families should participate in decision making.

No significant difference was found between managerial attitudes in relation to gender, work experience, academic background and *teaching branch* variables on a linear combination of communication, volunteering, learning at home, parenting, collaborating, and decision making. Managers' scores for learning at home, parenting and collaborating are lower than their scores for communicating, volunteering and decision making. The difference between the scores for their attitudes toward parental involvement dimensions may be due to the lack of parental practice. Therefore, there is a need to encourage more participation by families to enable them to gain more experience in school situations.

Teachers' attitudes toward building partnerships in communication, volunteering, learning at home, parenting, collaborating, and decision-making were not affected by their gender, work experience and *teaching branch*. On the other hand, teachers who had associate degrees expressed more positive attitudes toward involving parents through the communicating, volunteering, learning at home, and collaborating dimensions than those who had bachelor's or master's degree.

The findings showed that parents hold positive attitudes about effective parental involvement in Turkey. The struggle of the Ministry of National Education to involve parents in educational issues may be one of the most important reasons for these significant findings. The rate of female participants (344, or 56.6%) in undertaking the responsibility of the educational development of their children is another significant finding. The educational implication of this finding is that effective parent-school cooperation requires a revolutionary outlook: the focus must be on female parents.

It seems reasonable to suggest that encouraging experienced teachers to mentor other teachers about parental involvement practices may be a strategy to help increase parental involvement. Moreover, in Turkey, teacher training programmes do not currently include any significant systematized knowledge about parental involvement in ways outlined as potentially efficacious in the literature. As Greenwood and Hickman (1991) suggest, it is appropriate to require such a component not only in pre-service teacher training programmes, but also in in-service training. There are relevant techniques to apply and a number of books and materials available for that purpose. The following subjects for further empirical research are particularly suggested: (1) why do teachers and parents express lower rates than managers in all dimensions of building partnerships? (2) are there differences in mothers' and fathers' attitudes toward involvement at school? (3) what are the potential roles of other relatives in countries with traditional close-relative relations? (4) are there differences between the involvement of housewives and employed women, and in what ways can women contribute more? and finally, (5) future studies need to examine the whole area of potential parental involvement in secondary schools in Turkey.

Acknowledgements

This study was supported by the University of Akdeniz, the Department of Scientific Research Projects Management.

References

- Anderson, G. (1990). *fundamentals of educational research*. (London: The Farmer Press).
- Bagin, D., Gallagher, D. R. & Kindred, L. W. (1994). *The school and community relations* (5th edn.) Allyn and Bacon, Boston,
- Becker, H. J. & Epstein, J. L. (1982). Parental involvement: A survey of teacher practises, *Elementary School Journal*, 83, 85-102.
- Berger, H. P. (1991). Parental involvement: yesterday and today, *The Elementary School Journal*, 91(3), 209- 219.
- Bilgin, M. (1991). Ankara merkez ilçelerindeki ortaokullarda okul ile ailenin işbirliği ve sorunları [Parent-school cooperation and problems in secondary schools in central provinces in Ankara]. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, 45 (2), 89-116.

- Castro, D. C., Bryant, D. M., Peisner-Feinberg, E. S. & Skinner, M. L. (2004). Parental involvement in head start programs: the role of parental, teacher and classroom characteristics, *Early Childhood Research Quarterly*, 19, 413-430.
- Cervone, B. T., & O'Leary, K. (1982). A conceptual framework for parental involvement, *Educational Leadership*, 40 (2), 48-49.
- Cibulka, J. (1994). Policy options for building school-community relations, in: S. Lawton, E. Tanenzapf & R. Townsend (Eds.) *Education and Community: the Collaborative Solution Proceedings of the International Conference Linking Research and Practice* (University of Toronto, Ontario Institute for Studies in Education).
- Çelenk S. (2003). Okul aile işbirliği ile okuduğunu anlama başarısı arasındaki ilişki [*The relation between school-family cooperation and reading comprehension success*]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 24, 33-39.
- Çelenk, S. (2003a). Okul başarısının ön koşulu: okul aile dayanışması [*The prerequisite for school success: home-school cooperation*]. *İlköğretim-Online* 2 (2), 28-34.
- Dean, J. (2001). *Effective School Governor*. Florence KY, Routledge
- Çelik, N. (2005). *Okul-aile ilişkilerinde yaşanan sorunlar* [The Problems faced in school-parents relationship]. Unpublished Master Thesis. M.Ü. Atatürk Eğitim Fakültesi İlköğretim Anabilim Dalı, İstanbul.
- Dearing, E., McCartney, K., Weiss, H. B., Kreider, H. & Simpkins, S. (2004). The promotive effects of family educational involvement for low-income children's literacy, *Journal of School Psychology*, 42, 445-460.
- Demirbulak, D. (2000). Öğretmen görüşmeleri ile ilgili bir çalışma [A study about views of teachers]. *Milli Eğitim*, MEB Yayınlar Dairesi Başkanlığı, Ankara, 146.
- Epstein, J. L. (1984). School policy and parental involvement: Research result. *Educational Horizons*, 62 (2), 70-72.
- Epstein, J. L. (1987). Parental involvement: what research says to administrators. *Education and Urban Society*, 19 (2), 119-136.
- Epstein, J. L. (1995). School/family/community partnerships, *Phi Delta Kappan*, 76 (9), 701-713.
- Epstein, J. L. (2001). *School, family and community partnership. Preparing educators and improving schools*. Boulder, Colorado, West View Press.
- Epstein, J. L. & Dauber, S. L. (1991). School programs and teacher practices of parental involvement in inner-city elementary and middle schools, *The Elementary School Journal*, 91 (3), 289-305.
- Günkan H.E. (2007). *Ailenin ilköğretim öğrencilerinin eğitimi üzerindeki etki düzeyinin belirlenmesi* [Determination the efficiency level of family on education of primary school students] Unpublished Master Thesis. Fırat Üniversitesi Sosyal Bilimler Enstitüsü, Elazığ.

- Gürşimşek, I., Kefi, S., & Girgin, G. (2007). Okulöncesi eğitime babaların katılım düzeyi ile ilişkili değişkenlerin incelenmesi [Investigation of variables related with father involvement in early childhood education]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education)* 33, 181-191.
- Greenwood, G. E. & Hickman, C. W. (1991) Research and practice in parental involvement: implications for teacher education, *The Elementary School Journal*, 91(3), 279-288.
- Hair, J.F., Anderson, R.E., Tatham, R.L. & Black, W.C. (1998). *Multivariate data analysis*. Pentice Hall, Inc. Upper Saddle River: New Jersey.
- Henry, M. E. (1996). *Parental-school collaboration: feminist organizational structures and school leadership*. Albany NY, State University of New York Press.
- Hill, N. E. & Taylor, L. C., (2004). Parental school involvement and children's academic achievement, *Current Directions in Psychological Science*, 13 (4), 161-164.
- Dempsey, K. V. & Sandler, H. M. (2000). Why do parents become involved in their children's education?, in: L. Martyn (Ed) *Psychology of education Major themes: Schools, teachers and parents* Vol.1., Falmer Press, London.
- Jordon, C., Orozco, E. & Averett, A. (2001). *Emerging issues in school, family, and community connections*. National Center for Family & Community Connections with Schools/Southwest. Educational Development Laboratory, Austin.
- Kebeci, S. (2006). *Okul aile birliğinin okul performansı üzerindeki rolü* [The role of parent teacher association on school performance]. Unpublished Master Thesis. Marmara Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı Yönetim ve Organizasyon Bilim Dalı, İstanbul.
- Lasky, S. (2000). The cultural and emotional politics of teacher-parental interactions, *Teaching and Teacher Education*, 16, 843-860.
- Martinez, R. A., Martinez R. & Perez, M .H. (2004). Children's school assessment: Implications for family-school partnerships. *International Journal of Educational Research*. 41, 24-39
- M.E.B. (2005). *Öğrenci-veli-okul sözleşmesi, eğitimi araştırma ve geliştirme dairesi başkanlığı genelgesi*, Sayı: 2005/92.
- Miedel, W. T. & Reynolds, A. J. (1999). Parental involvement in early intervention for disadvantaged children: does it matter? *Journal of School Psychology*, 37(4), 379-402.
- Muijs, D. (2004). *Doing quantitative research in education with spss*. Sage Publications Ltd., London.
- Resmi Gazete. (1961). *İlköğretim ve eğitim kanunu*, Sayı: 10705.
- Resmi Gazete. (2003). *Milli eğitim bakanlığı ilköğretim kurumları yönetmeliği*, Sayı: 25212.
- Resmi Gazete. (2005a). *Milli eğitim bakanlığı okul-aile birliği yönetmeliği*, Sayı:25831.

- Resmi Gazete, (2005b). *Milli eğitim bakanlığı ilköğretim ve ortaöğretim kurumları sosyal etkinlikler yönetmeliği*, Sayı: 25699.
- Voorhis, F. V. & Sheldon, S. (2004). Principal's roles in the development of us programs of school, family, and community partnerships, *International Journal of Educational Research*, 41, 55-70.
- Werf, G. V. D., Creemers, B. & Guldemon, H. (2001). Improving parental involvement in primary education in indonesia: Implementation, effects and costs, *School Effectiveness and School Improvement*, 12 (4), 447-466.
- Williams, D. L., & Chavkin, N. F. (1989). Essential elements of strong parental involvement programs. *Educational Leadership*, 47 (2), 18-20.
- Yıldırım, M. C. & Dönmez B. (2008). Okul aile işbirliğine ilişkin bir araştırma: istiklal ilköğretim okulu örneği. [A study on the cooperation between the school and parents (İstiklal primary school sample)] *Elektronik Sosyal Bilimler Dergisi*, Kış-C.7. (23), 98-115.
- Yılmaz, (2006). *İlköğretim okullarında etkililik sağlanması için öğretmenlerin ailelerden beklentileri (Sakarya ili örneği)* [The expectations of the teachers from the families in order to provide effectiveness in elementary education schools]. Unpublished Master Thesis. Sakarya Üniversitesi. Sosyal Bilimler Enstitüsü, Sakarya.
- Zellman, G. L. & Waterman, J. M. (1998). Understanding the impact of parental school involvement on children's outcomes, *The Journal of Educational Research*, 91(6), 370-380.

Türkiye’de Veli Katılımına İlişkin İlköğretim Okulu Yöneticilerinin, Öğretmenlerinin ve Öğrenci Velilerinin Görüşleri

(Özet)

Problem Durumu: Öğretmenlerin ve öğrencilerin etkileşim şekillerinin etkili bir öğrenme ve eğitim çevresi oluşturmada önemli bir etkisi vardır. Aile pek çok etkenin yanında okula hem sınıf içi öğretim süreçlerinde hem de okulun tüm yönetim süreçlerinde gönüllü katkı verebilecek önemli bir kaynak olarak değerlendirilmektedir. Geleneksel bakış açısına göre aileler, okulun başat bir parçası olmaktan çok temsilci veya destekleyici bir faktör olarak değerlendirilmiştir. Öğretmenler eğitimi ve öğretimi gerçekleştirmede tek ve yeterli uzman olarak görülmüştür. Çocuğun öğrenme gereksinimlerini karşılama konusunda aile göz ardı edilmiş öğretmen tek başvuru kaynağı olarak görülmüştür. Çocuğun eğitim ihtiyacının karşılanmasında yetki eğitimcilerle verilmiştir. Aile okulun verdiği görevleri yerine getiren bir öge olarak değerlendirilmiştir. Ancak son yıllarda okulun amaçlarını aileler veya veliler olmadan gerçekleştiremeyeceği anlaşılmıştır. Türkiye’de de ailelerin okulun tüm süreçlerine etkin katılımı için yasal ve yönetsel düzenlemeler yapma yönünde önemli çabalar görülmektedir.

Araştırmanın amacı: Bu araştırmanın amacı öncelikle yöneticilerin, öğretmenlerin ve ailelerin okul ile aile arasında daha etkili bir işbirliği oluşturmaya ilişkin düşüncelerini belirlemek ve ikinci olarak da bu tutumlarının meslek, eğitim durumu, cinsiyet, kıdem, brans, ailelerin meslekleri, gelir düzeyleri, çocuğunun okuduğu sınıf düzeyi gibi çeşitli faktörlere göre farklılaşıp farklılaşmadığını ortaya koymaktır. Elde edilen veriler bir yandan uygulama bakımından Türkiye’de okul-aile işbirliğinin geliştirilebilmesine katkı sağlamak bir yandan da bu alanda çalışma konusunda isteklilik gösteren araştırmacılara ışık tutmak bakımından önem taşımaktadır.

Araştırmanın Yöntemi: Araştırma alan taraması modelinde desenlenmiştir. Araştırmanın evreni Antalya ili merkez ilçelerde bulunan 126 devlet ilköğretim okulunda görev yapan okul müdürleri, 199 müdür yardımcısı, 4425 öğretmen ve 7 eğitim bölgesindeki ilköğretim okulu öğrencileri ve velilerinden oluşmuştur. Araştırmada örneklem ise küme örnekleme yöntemi ile belirlenen 35 ilköğretim okulunda bulunan yöneticiler, 277 öğretmen, belirlenen 35 okuldaki tüm sınıfların öğrencileri arasında tesadüfî yöntemle seçilmiş 840 öğrencinin velisinden oluşmuştur. Araştırmada, araştırmacı tarafından geliştirilmiş olan bir ölçek kullanılmıştır. Ölçek, Likert tipi, biri yöneticilere ve öğretmenlere; biri de velilere uygulanmak üzere hazırlanmış iki formdan oluşmaktadır. Ölçek, iletişim, gönüllü katılım, evde öğrenme, destekleyici ev ortamı, işbirliği ve karar verme olmak üzere 5 boyuttan oluşmaktadır. Ölçeğin yönetici ve öğretmen formunun güvenilirlik katsayısı toplamda (Cronbach Alpha).95; velilere uygulanan ölçeğin ise güvenilirlik katsayısı toplamda (Cronbach Alpha).91’dir. Verilerin analizinde SPSS paket programından yararlanılmış ve frekans, aritmetik ortalama, standart sapma ve parametrik testlerden MANOVA ve one way ANOVA kullanılmıştır.

Araştırmanın Bulguları: Yöneticilerin, öğretmenlerin ve velilerin etkili okul-aile işbirliğine ilişkin bütünde olumlu tutuma sahip oldukları anlaşılmıştır. Ancak yöneticilerin diğer gruplara göre daha yüksek puanlar verdikleri bulunmuştur. Yöneticilerin görüşleri arasında cinsiyete, kıdeme, eğitim durumuna ve bransa göre bir farklılık bulunmamıştır. Velilerin görüşlerinde de cinsiyet'e, velinin mesleğine, çocuğunun sınıf düzeyine, velinin gelir durumuna ve velinin eğitim durumuna göre fark bulunmamıştır. Ön lisans düzeyinde öğrenim gören öğretmenlerin görüşleri ile lisans ve lisansüstü düzeyde öğrenim gören öğretmenlerin görüşleri arasında iletişim, evde öğrenme ve işbirliği boyutlarında anlamlı farklar belirlenmiştir.

Araştırmanın Sonuçları ve Önerileri: Araştırmaya katılanların demografik özellikleri değerlendirildiğinde kadın öğretmen sayısının oransal olarak yönetimde yeterince temsil edilmediğini göstermektedir. Öte yandan yöneticilerin önemli bir bölümünün de ön lisans eğitimi almış oldukları görülmüştür. Her iki durumun da okul içi etkili çalışma ilişkileri açısından önemli engeller oluşturduğu söylenebilir. Ortaya çıkması olası bu durumun da okul- aile ilişkilerine olumsuz yansımaları söz konusu olabilir. Ailelerin okulla etkili işbirliği yapmasına ilişkin yönetici, öğretmen ve velilerin tutumları genel olarak olumlu olmakla birlikte kendi içlerinde farklılıklar göstermektedir. Buna göre öğretmenler ile velilerin işbirliğine ilişkin tutumlarına ilişkin puanları daha düşüktür. Buna göre, öğretmenlerin ve velilerin, veli katılımına ilişkin tutumları ile ilgili yasal ve yönetsel çerçevede farklı yaklaşımlar geliştirilmesi gerektiği söylenebilir. Yöneticilerin evde öğrenme, destekleyici ev ortamı ve işbirliği boyutlarında verdikleri puanların iletişim, gönüllü katılım ve karar verme boyutlarına verdikleri puanlardan daha düşük olması mevcut durumda da sorunlu görünen üç alan ile ilgili yeterli deneyimin bulunmaması ile ilişkilendirilebilir. Bu durumda özellikle evde öğrenme, destekleyici ev ortamı ve işbirliği boyutları ile ilgili üzerinde uzlaşmış, tarafların gereksinimleri ile örtüşen yaratıcı deneyimlere gereksinim olduğu söylenebilir. Ailelerin de olumlu tutuma sahip olmaları anlamlı bir sonuçtur. Çünkü mevcut yapıda öğretmenlerle ve yöneticilerle ilişkileri kendilerinden taleplerden oluşan ve kendilerine çoğunlukla tek yönlü iletiler gönderilerek olumsuz duygular geliştirmelerine neden olan ailelerin bu işbirliğine yatkın olmaları önemlidir. Bu konuda kritik nokta ekonomik, sosyal ve kültürel durumlarının; ilgi ve yeteneklerini dikkate alan yeni bir işbirliğine zemin hazırlama konusunda bilinçli çabaların sergilenmesidir. Bütünde değerlendirildiğinde etkili okul aile işbirliğinin kurulabilmesi için tarafların sahip oldukları olumlu tutum etkili bir başlangıç için önemli bir veri oluşturmaktadır. Bununla birlikte Bu olumlu tutumun pratik uygulamalar bakımından da alışlagelmişin dışında okul aile işbirliğinin önündeki mevcut engelleri dikkate alan devrimsel nitelikli yaratıcı uygulamalar gerektirdiği söylenebilir. Bu bakımdan tüm tarafların katılımı ile ve tüm tarafların sosyal, ekonomik, kültürel gereksinimleri ile örtüşen yeni yöntem ve stratejiler geliştirilmesi gerekmektedir.

Anahtar Sözcükler: okul-aile işbirliği, veli katılımı, öğrenci başarısı, etkili okul

A Study on Developing a Self Efficacy Scale towards Science and Technology*

Nilgün Tatar**
Eylem Yıldız***
Ercan Akpınar****
Ömer Ergin*****

Suggested Citation:

Tatar, N., Yıldız, E., Akpınar, E. & Ergin, Ö. (2009). A Study on Developing a Self-Efficacy Scale towards Science and Technology. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 36, 263-280.

Abstract

Problem Statement: On examining the renewed elementary education program, it was seen that the technology dimension has been added to elementary education science lessons. When the related literature was reviewed, it was observed that there were no measurement tools to determine the self-efficacy of elementary school students regarding science and technology in Turkey. Therefore, there is a need for a measurement tool which can establish students' self-efficacy in science and technology education.

Purpose of Study: The aim of this study was to develop the "Science and Technology Self-Sufficiency Scale" (SESST) for the purpose of evaluating the self-efficacy in science and technology of elementary school students and to conduct a validity and reliability assessment of this scale and thus provide the scale for use in the field of education.

Methods: The sample consisted of 10 primary schools with 400 students in grades 6, 7 and 8. In order to examine the construct validity of the SESST, an exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were performed. To ensure the validity of the SESST, corrected item-total correlations and t-tests between items' means of upper 27% and lower 27% points and Croanbach alpha correlation coefficients were used. Additionally, the means and standard deviations for the factors were examined; Pearson Product Moment Correlation Coefficients were calculated among these factors.

Findings and Results: The EFA and DFA results showed that the factorial model of the scales consisted of three factors. To establish that each scale

* This study was supported by TUBITAK (Project # 106K268)

**Asst. Prof. Dr., Cumhuriyet University, ntatar@cumhuriyet.edu.tr

*** Dr., Cumhuriyet University, eylemyildiz@cumhuriyet.edu.tr

**** Lecturer,Dr., Dokuz Eylül University, ercan.akpinar@deu.edu.tr

***** Prof. Dr.,Dokuz Eylül University, omer.ergin@deu.edu.tr

had satisfactory internal consistency, Cronbach's alpha coefficient was calculated. The internal consistency of "Confidence in Science and Technology Ability" was 0.93, "Coping with Difficulties in Science and Technology" was 0.75, and "Confidence in Performing Science and Technology Tasks" was 0.80. Each scale displayed satisfactory internal consistency reliability. The overall scale reliability was calculated as 0.93.

Conclusions and Recommendations: The SESST may be a useful tool in identifying the self-efficacy of elementary school students in science. The scale presented in this study provides an instrument that can be used with students in elementary and middle school classrooms as well. Prospective studies may examine the relationships between the SESST and other variables, i.e. academic performance, development level of students, and gender. The scale, which has been tested for validity and reliability, is considered useful for research into science and technology training.

Keywords: Elementary school students, science and technology education, self-efficacy, scale development

Self-efficacy is defined as a person's belief in his or her ability to acquire new information or to complete a task or activity to a prescribed level of performance. Bandura (1977, 1986) put forward the concept of self-efficacy in explaining human beings' behaviors, together with social learning theory. According to him, this concept expresses the self-belief of an individual in accomplishing a specific task successfully. In other words, this concept is an expression of thoughts, behaviors, beliefs and feelings of individuals on any matter. According to Senemoğlu (1997), self-efficacy is an outcome of judgments related to an individual's capabilities. Kurbanoglu (2004) suggests that self-efficacy, as one of the most important variables that affects human behaviors, reflects a person's beliefs in his/her own abilities to accomplish a mission.

According to Bandura (1994), when faced with obstacles and failures, people who harbor self-doubts about their capabilities slacken their efforts or give up quickly. Those who have a strong belief in their capabilities exert greater efforts even when they fail to master the challenge. Strong perseverance contributes to performance accomplishments. Pajares (2000) also argued that when students lack confidence in their capabilities, they are likely to attribute their failure to low ability, which they perceive as inborn, permanent, and not acquirable. For them, failure is just another reminder that they are incapable. Those who have a high self-efficacy perception behave more like fighters and undertake responsibilities. Thus, it can be said that self-efficacy affects the individual in the accomplishment of a task (Bandura, 1977).

The self-efficacy concept, which Bandura claims affects the behavior of an individual, is an individual's judgment about how he/she will be successful in coping with the problems which he/she might encounter in the future. People with high self-efficacy are capable of making decisions, setting goals and displaying commitment. Achieving independently and within groups was similar (Singer, 2008). The self-efficacy belief of an individual affects his/her motivation and performance

in different forms. Self-efficacy is, in addition, closely related to attitudes and behaviors (Morgil, Seçken & Yücel, 2004).

Self-efficacy scales have been developed for different fields (education, psychology, health and sports) in order to show that it has a strong influence on an individual's characteristics. The scales in educational fields are used to identify students' and teachers' self-efficacy levels. Studies on self-efficacy used to explain the behaviors of teachers and students in education literature have been conducted by employing various scales developed since the late 1970s. Many studies have been conducted on the concept of self-efficacy, which was originally put forward by Bandura, in every area of education, and relevant scales have been developed (Table 1).

Since Bandura, self-efficacy theory has been applied in educational settings within different grade levels, i.e. elementary, primary and secondary (Schunk, 1996). Various researchers have studied and reported on the relationships between self-efficacy and academic performance. Kupermintz (2002) argues the importance of considering motivational resources in explaining the performance of academic tasks. In Kupermintz's study, students who reported the high value of and a strong interest in science as a domain were more likely to achieve high scores on a multiple-choice test, especially within the basic knowledge and spatial-mechanical dimensions, than to receive higher science grades for their performance in class. Multon, Brown and Lent (1991) found 31 studies written between 1977 and 1988 on the relationships between self-efficacy and academic performance (cited in Adeyemo, 2007). In the results of the study, it was found that efficacy beliefs were related to performance ($r=.38$). The studies on teacher candidates show that as students succeed in science courses, their level of self-efficacy belief increases (Yavuzer & Koç, 2002; Ören-Şaşmaz & Tatar, 2006).

Furthermore, in research conducted at a college level (Rose, 2003), the results of the study showed highly significant positive relationships between self-efficacy and academic achievement. Based on the research, it can be said that there is a correlation between self-efficacy and academic success.

There has also been research investigating self-efficacy at the classroom level using gender, self-concept and self-regulation variables. In the research by Umay (2002) in which she analyzed the effects of a Primary School Mathematics Teaching Program on a self-efficacy perception of mathematics, it was shown that the last grade students' self-efficacy perceptions were significantly higher than those of the first graders. Ören-Şaşmaz and Tatar's collective research (2006) on primary school teacher candidates' self-efficacy beliefs in relation to science courses showed that there were no significant differences between groups in relation to gender. Additionally, it has been observed that self-efficacy beliefs increase as the grade level increases.

Table 1
Some Scales Regarding Self-Efficacy

Scale	Researchers	Sampling
➤ The Self-efficacy Belief Scale in Science Teaching for Primary School Teachers	Enochs and Riggs (1990) (Adapted by Hazır-Bıkmaz, 2004)	Classroom teachers
➤ Science Teaching Self-Efficacy Scale	Gibson and Dembo (1984)	Elementary school teachers
➤ Self-Efficacy Perception Scale Relevant to Computer	Aşkar and Umay (2001)	Pre-service elementary school mathematics teachers
➤ The Teacher Interpersonal Self-Efficacy Scale	Brouwers and Tomic (2002) (Adapted by Çapri and Kan, 2007)	Teachers working for the primary, secondary, private school and preparatory courses
➤ Self-Efficacy Scale Towards Geometry	Cantürk-Günhan and Başer (2007)	Elementary school students
➤ Science Self-Efficacy Scale	Lian (2003)	Secondary school students
➤ Self-Efficacy Belief Scale Intended for Chemistry Teaching	(Morgil et al., 2004)	Pre-service chemistry teachers
➤ Science Teaching Professional Self-efficacy Inventory	Riggs and Enoch (1990) (Adapted by Önen and Öztuna, 2005)	Science and mathematics teachers
➤ The Self-efficacy Scale for Computer Teaching	Akkoyunlu, Orhan and Umay (2005)	Teacher candidates for computer and teaching technology training

Vural and Hamurcu (2008) examined preschool teacher candidates' self-efficacy beliefs regarding the teaching of science lessons. The results of the study showed that the third class preschool teacher candidates' self-efficacy beliefs and outcome expectation points were higher compared to the first class preschool teacher candidates, and there were significant differences between self-efficacy points in favor of the third class preschool teacher candidates. Choi (2005) examined the relationship between self-efficacy and self-concept measured at varying degrees of specificity using the academic performance of college students. The findings of the study showed that specific self-efficacy was the only significant predictor of term grades. Otherwise, college students who had a high degree of self-perceptions tended to attain higher levels of academic achievement. According to Choi, one implication that can be drawn from this finding relates to the importance of designing classroom activities that would aid in enhancing college students' self-concept and self-efficacy. Nevill (2008) examined the relationship between reading self-efficacy and the regulation of cognition, important components in the development of self-regulated learning, and reading achievement in elementary school students. According to the results, reading self-efficacy is a predictor of both the regulation of cognition and reading achievement in an intermediate elementary sample. The finding supports the premise that students are more efficacious about their ability to read and tend to regulate their cognition at a level significantly different from those with lower reading self-efficacy. While positively associated

with reading achievement, a significant relationship between the regulation of cognition and reading self-efficacy does not exist.

As a result of the literature review, it was discovered that a scale prepared for measuring the self-efficacy of elementary school students of science and technology does not exist in Turkey. Since there are no scales to measure students' self-efficacy in science and technology courses, there is a need to develop a scale in this field. The purpose of this study was to develop a scale to assess students' self-efficacy regarding science and technology and to examine the validity and reliability of this scale.

Method

Sample

The sample consisted of 10 primary schools with 400 students in grades 6, 7 and 8. Before administering the questionnaire to all participants, a pilot study was conducted to check the readability and comprehensibility of the questionnaire items by interviewing six students, four science teachers and three scientists. By taking the students', teachers' and scientists' feedback into consideration, some minor revisions and modifications were made. Then, one of the researchers visited each of the 10 schools and explained the purpose of the questionnaire, read the instructions aloud and answered any individual questions that the students asked. Table 2 shows the students' grade levels and characteristics according to their gender.

Table 2

The Demographic Characteristics for the Students Who Were Given a Scale Test

		Gender (N)		
		Male	Female	Total
Grade level	6. Grade	60	68	128
	7. Grade	78	62	140
	8. Grade	63	69	132
	Total	201	199	400

In the first stage of the current study, the scale, "Self-Efficacy for Science and Technology" (SESST) (Appendix A) consisting of 36 items, was developed by the researchers by revising and referring to some studies regarding other self-efficacy scales (Chen, Gully & Eden, 2001; Haskell, 2002; Işıksal & Aşkar, 2003; Thompson, 2003; Bleicher, 2004; Hazır-Bıkmaz, 2004; Akkoyunlu, et al., 2005; Hampton, 2005; Kurbanoglu, Akkoyunlu, & Umay, 2006). Once the scale was created, views of three science and technology teachers and three professors were sought for the face validity of the scale; required changes were made based on the feedback received. A Likert-type response format was provided with response options ranging from (1) strongly disagree to (5) strongly agree.

Data Analysis

All statistical analyses were performed using SPSS statistical software 10.00 and LISREL 8.80. The exploratory factor analysis and confirmatory factor analysis were performed to determine the construct validity of the scale. The principal components factor analysis was followed by varimax rotation (rotated component matrix). We used varimax rotation because the final factors would be as uncorrelated as possible with each other. We thought that the variance explained by one factor would be independent of the variance in the other factors. Rotation makes it so that, as much as possible, different items are explained or predicted by different underlying factors, and each factor explains more than one item. This is a condition called simple structure (Leech, Barrett & Morgan, 2005).

For the validity of the SESST, corrected item-total correlations and t-tests between items' means of upper 27% and lower 27% points and Croanbach alpha correlation coefficients were used. Besides this, the means and standard deviations for the factors were examined; Pearson Product Moment Correlation Coefficients were calculated among these factors.

Findings

Exploratory Factor Analysis

Before conducting the factor analysis of the scale, the Kaiser-Meyer Olkin (KMO) measure of sampling adequacy (KMO) and Barlett's test were calculated to evaluate whether the sample was large enough to perform a satisfactory factor analysis. The KMO sampling adequacy test statistic was 0.95. This value is higher than the threshold value of 0.5 (Hair, et al., 1998). Barlett's test of Sphericity statistic was significant [6616.83 ($p < 0.01$)]. Thus, these results appear to support the validity of the factor analysis usage for this study.

The purpose of the exploratory factor analysis was to investigate the factors underlying the SESST. The exploratory factor analysis was performed on the 36 items. The initial solution revealed that six factors had an eigenvalue greater than 1, explaining 55.98% of the variance. After using varimax rotation, the factor loadings for each item were examined. Loadings of less than 0.30, a commonly-used cut-off, were eliminated. It was observed that one factor had only two items. Besides this, six items had a high loading on more than one factor. Therefore, these eight items were eliminated. Then, varimax rotation was maintained and replicated twice. The final solution had three factors with an eigenvalue greater than 1.

Results of the factor extraction are presented in Table 3. Table 3 shows the results of the factor analysis for the 27-item SESST for the sample of 400 students. All three factors explain 51% of the variance.

Table 3
Factor Loading of Items in the SESST

Item	Factor loading		
	Factor 1	Factor 2	Factor 3
35	.775		
29	.750		
33	.724		
23	.717		
30	.691		
28	.681		
20	.679		
32	.665		.385
36	.661		
22	.628		.332
34	.628		
31	.628	.313	
25	.628	.376	
18	.604		
15	.549		
2		.680	
8		.674	
6		.633	
24		.616	
16		.606	
17		.451	
9			.732
11			.645
21	.314		.644
13	.305	.312	.639
26		.429	.560
27	.301		.464
Eigenvalue	10.672	1.917	1.320
% Variance	39.526	7.100	4.890

*Factor loadings smaller than 0.30 have been omitted.

Confirmatory Factor Analysis

Through a LISREL confirmatory factor analysis, the items for each scale of efficacy were examined for convergent validity and construct validity. The fit of the model was evaluated with various measures (Sörbom & Jöreskog, 1982). Kelloway (1998) has suggested that the use of the chi-square test is reasonable when the study involves a large sample. However, as the chi-square is very sensitive to sample size, the degree of freedom can be used as an adjusting standard by which to judge whether chi-square is large or small (Jöreskog, & Sörbom, 1989). Other types of goodness-of-fit measures include Root Mean Squared Error of Approximation (RMSEA), Root Mean Square Residual (RMR), Standardized RMR (SRMR), normed

fit index (NFI), non-normed fit index (NNFI), the comparative fit index (CFI), Goodness of Fit Index (GFI), and the Adjusted Goodness of Fit Index (AGFI). RMSEA and RMR values close to zero show a near perfect fit. The NFI, NNFI, CFI, GFI, AGFI are always between zero and one, with any value above 0.9 indicating a good fit and the value one suggesting a perfect fit. Results of goodness-of-fit measurements were: $\chi^2 = 619.23$ (df=321, $p < 0.000$), (χ^2/df)=1,92, Goodness of Fit Index (GFI)=0.89, Normed Fit Index (NFI)= 0.97, Non-Normed Fit Index (NNFI) = 0.99, Root Mean Square Residual (RMSEA)=0.048, Adjusted Goodness of Fit Index (AGFI)= 0.87, Comparative Fit Index (CFI)= 0.98, Root Mean Square Residual (RMR)= 0.06 and Standardized RMR (SRMR) = 0.041. All of the t-values of items showed statistical significance at the 0.05 level. The ratio of the model was 1.92, indicating a fairly good fit. The model had a RMSEA of 0.0048, SRMR of 0.041, and NFI, NNFI, and CFI values over 0.9, showing that the model had a highly satisfactory fit. Results also indicated that there was an appreciable relationship between the error covariance of items 12 and 13 as well as 20 and 21. In this context, it was decided to test the error correlations, which were observed among the items that exist under the same factor (latent variable) in the scale, by adding them to the model and performing CFA again. Types of goodness-of-fit measures for the second CFA were: $\chi^2 = 519.85$ (df=317), $p < 0.00000$, (χ^2/df) = 1.63, RMSEA=0.041, GFI=0.91, NFI=0.97, NNFI=0.99, RMSEA=0.042, AGFI=0.89, CFI=0.99, RMR=0.060, SRMR=0.042. These values showed that the model had a highly satisfactory fit.

Subsequent to the analyses, the items in the factors were examined, and the factors were named. The literature research was made use of when naming these factors (Chen et al., 2001; Mangos & Steele-Johnson, 2001; Britner & Pajares, 2006). Scale description and sample test items are reported in Table 4.

Table 4
Scale, Description and Sample Test Items from the Scale

Scale	Description	Sample Item
• Confidence in science and technology ability	Extent to which student has confidence in science technology ability.	I am sure that I can accomplish all skills given in a science and technology lesson successfully.
• Coping with difficulties in science and technology	Extent to which student copes with difficulties in science and technology.	I am not sure that I can understand difficult scientific and technologic concepts.
• Confidence in performing science and technology tasks	Extent to which student has confidence in performing science and technology tasks.	If scientific activities are hard, I give up or I only accomplish parts that are easier.

Item Analysis

After the confirmatory factor analysis, the differences between mean scores of upper 27% and lower 27% points were examined for each item. The t-test results showed significant differences between each item's means of upper 27% and lower 27% points. All corrected item-total correlations ranged from 0.33 to 0.69. Results are presented in Table 5.

Table 5

SESST Factors, Regulated Items, Total Correlations, Item-Total Correlations and Upper 27 %, Lower 27 % Points, Unrelated t-Test Results

Number of Item	Corrected Item-Total Correlation	t values for items (Upper 27%-Lower 27%)	Number of Item	Corrected Item-Total Correlation	t values for items (Upper 27%-Lower 27%)
Confidence in science and technology ability			Coping with difficulties in science and technology		
35	.6871	14,827	2	.3386	7,93
29	.6476	11,760	8	.5246	11,555
33	.6855	17,896	6	.4729	9,775
23	.6947	15,718	24	.4439	10,452
30	.6562	14,378	16	.4945	12,637
28	.6455	14,674	17	.4459	10,365
20	.6138	12,590	Confidence in performing science and technology tasks		
32	.6834	13,395	9	.4425	9,514
36	.6223	11,574	11	.4971	10,247
22	.6288	11,767	21	.5809	11,162
34	.5741	10,240	13	.6334	11,715
31	.6726	14,715	26	.6025	13,497
25	.6950	15,916	27	.5200	12,873
18	.6061	13,821			
15	.5711	11,829			

The average and standard deviation related to the scale's three identified factors are given in Table 6. When the binary correlations between factor points are examined, the correlation multipliers are observed as positive and meaningful ($p < 0.01$).

Table 6

The Average and Standard Deviation of the SESST Factors and Inter-Factorial Correlation Values

		\bar{X}	SS	Correlations*		
Factors				1	2	3
1	Confidence in science and technology ability	30,46	12,05	-	,546	,659
2	Coping with difficulties in science and technology	16,00	5,48			,566
3	Confidence in performing science and technology tasks	11,44	5,32			-

* Correlation is significant at the 0.01 level.

Conclusions and Implications

The purpose of this study was to develop the reliability and validity of the SESST for students in elementary school in general. When developing a questionnaire,

content validity and construct validity should be identified. For content validity, we used previous studies from different settings, different students' levels of self-efficacy and also used existing questionnaires to design the questionnaire items. In addition to three science teachers, three science educators and eight elementary school students reviewed all the test items. Construct validity was verified by a factor analysis.

We analyzed a set of 32 items. The questionnaire consisted of three factors and 27 items. The internal consistency of "Confidence in Science and Technology Ability" was 0.93, "Coping with Difficulties in Science and Technology" was 0.75, and "Confidence in Performing Science and Technology Tasks" was 0.80. The overall scale reliability was calculated as 0.93. According to the results, we have identified three important factors in exploring students' self-efficacy: students' confidence in their science and technology abilities, coping with difficulties in science and technology, and confidence in performing science and technology tasks.

Students' trust in their abilities to succeed in science and technology courses (mental, perceptual or physical) may affect their self-efficacies towards these courses. The confidence that students have in their own abilities may help to determine what they do with the knowledge and skills they possess. Pajares and Kranzler (1995) stated that ability and self-efficacy had strong direct effects on performance. Ability also has a strong direct effect on self-efficacy, which largely mediates the indirect effect of ability and background on performance. Also, reasoning ability influences self-efficacy. Presumably, this is because reasoning ability plays a role in determining whether or not a person is successful in completing the sorts of tasks used in the present study, and prior successes increase self-efficacy (Lawson, Banks & Logvin, 2007). Bandura (1997) believed that people with high cognitive self-efficacy were determined to believe in their abilities, to plan strategies, anticipate the probable results and to set goals (cited in Singer, 2008). Self-efficacy researchers posit that students' beliefs in their abilities to succeed in science tasks, courses, or activities, or their science self-efficacy, influence their choices of science-related activities, the effort they expend on those activities, the perseverance they show when encountering difficulties and the ultimate success they experience in science (Britner & Pajares, 2006).

Facing difficulties, fighting against them and getting over them can improve an individual's self-efficacy. Self-efficacy beliefs have been tied to increased levels of task persistence (Wolters, 2003). Being firmly resolute in attaining goals through adjustment based on steady progress causes individuals to be highly motivated. Instead of concentrating on personal imperfections, these individuals think positively about their aspirations, prepare for difficult challenges, and dedicate time to facing their difficult challenges (Singer, 2008). Mangos and Johnson (2001) posited that an individual's perception of task complexity would be influenced by the individual's cognitive ability and self-assessments of competence (i.e., self-efficacy) to perform the task. However, subjective task complexity would contribute to the prediction of task performance beyond the variance explained by ability or self-efficacy. Chen et al. (2001) indicated that cognitive ability and conscientiousness positively relate to self-efficacy, but that the magnitude of these relationships varies with task complexity. In

science, students who have a strong belief that they can succeed in science tasks and activities will be more likely to select such tasks and activities, work hard to complete them successfully, persevere in the face of difficulty and be guided by physiological indexes that promote confidence as they meet obstacles. Alternatively, students who do not believe that they can succeed in science-related activities will avoid them if they can and will make a minimal effort if they cannot. When confronted with the typical challenges that science involves, they will be more likely to give up and to experience the stresses and anxieties that help ensure the erosion of their efforts (Britner & Pajares, 2006).

An individual's self-efficacy affects the fulfillment of a mission assigned to him/her. People with firm assurance in their capabilities approach difficult tasks as challenges to be mastered rather than as threats to be avoided. Such an efficacious outlook fosters intrinsic interest and deep engrossment in activities. These kinds of people set challenging goals for themselves and maintain strong commitment to them. They heighten and sustain their efforts in the face of failure. They quickly recover their sense of efficacy after failures or setbacks. They attribute failure to insufficient effort or deficient knowledge and skills, which are acquirable. They approach threatening situations with assurance that they can exercise control over them. Such an outlook produces personal accomplishments, reduces stress and lowers vulnerability to depression (Bandura, 1994). Eccles-Parsons et al. (1983) defined expectancy as individuals' beliefs about how well they would perform future tasks in a given domain. These theories posit that expectancies for success and values serve the function of preparing and energizing individuals to engage in a task, to seek out task challenges and to persist at particular tasks, all contributing to successful performance via a process Snow called "the commitment pathway" (cited in Kupermintz, 2002).

In conclusion we believe that the SESST may be a useful tool for identifying the self-efficacy of primary students in science. The SESST has previously been used only with university students in Turkey; the scale presented in this study provides an instrument that can be used with students in elementary and middle school classrooms as well. Prospective studies may examine the relationships between the SESST and other variables, i.e. academic performance, development level of students and gender. The scale, which tested for validity and reliability, is considered useful for the research into science and technology training.

References

- Adeyemo, D.A. (2007). Moderating influence of emotional intelligence on the link between academic self-efficacy and achievement of university students. *Psychology and Developing Societies*, 19 (2), 199-213.
- Akkoyunlu, B., Orhan, F., & Umay, A. (2005). Bilgisayar öğretmenleri için "bilgisayar öğretmenliği öz yeterlik ölçeği" geliştirme çalışması [A study on developing teacher self efficacy scale for computer teachers]. *Hacettepe Üniversitesi, Eğitim Fakültesi Dergisi*, 29, 1-8.

- Aşkar P. & Umay A. (2001). İlköğretim matematik öğretmenliği öğrencilerinin bilgisayarla ilgili özyeterlik algısı [The effect of the primary school mathematics teaching program on the mathematics self-efficacy of students]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 21, 1-8.
- Bandura, A. (1977). Self- efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84 (2), 191-215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1994). *Self-efficacy*. V.S. Ramachaudran (Ed.), *Encyclopedia of human behavior* Vol. 4. (pp. 71-81), New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. San Diego: Academic Press, 1998). Retrieved July 10 2008 from <http://www.des.emory.edu/mfp/BanEncy.html>
- Bleicher, R. E. (2004). Revisiting the STEBI-B: measuring self-efficacy in preservice elementary teachers. *School Science and Mathematics Journal*, 104, 383-391.
- Britner, S. L. & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, 43 (5), 485-499.
- Cantürk-Günhan, B. & Başer, N. (2007). Geometriye yönelik öz-yeterlik ölçeğinin geliştirilmesi [The development of self-efficacy scale toward geometry]. *Hacettepe Üniversitesi, Eğitim Fakültesi Dergisi*, 33, 68-76.
- Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, 4 (1), 62-83.
- Chen, G., Casper, W. J., & Cortina, J. M. (2001). The roles of self-efficacy and task complexity in the relationships among cognitive ability, conscientiousness, and work-related performance: A meta-analytic examination. *Human Performance*, 14 (3), 209-230.
- Choi, N. (2005). Self-efficacy and self-concept as predictors of college students' academic performance. *Psychology in the Schools*, 42 (2), 197-205.
- Çapri, B. & Kan, A. (2007). Öğretmenlerin kişilerarası öz-yeterlik inançlarının hizmet süresi, okul türü, öğretim kademesi ve unvan değişkenleri açısından incelenmesi [The investigation of the teachers' interpersonal self-efficacy beliefs according to working experiences, type of school worked in, educational level worked in, and professional position]. *Mersin University Journal of the Faculty of Education*, 3 (1), 63-83.
- Ekici, G. (2005). Biyoloji öz- yeterlik ölçeğinin geçerlik ve güvenilirliği [The validity and reability of the biology self- efficacy instrument]. *Hacettepe University Journal of Education*, 29, 85-94.
- Gibson, S. & Dembo, M. H. (1984), Teacher efficacy: A construct validation. *Journal of Educational Psychology*, 76 (4), 569-582.
- Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C., (1998). *Multivariate data analysis* (5th edition), New Jersey, Prentice Hall.
- Hampton, N. Z. (2005). Testing for the structure of the career decision self-efficacy scale-short form among Chinese college students. *Journal of Career Assessment*, 13 (1), 98-113.

- Haskell, D. H. (2002). *Lecture to inquiry: The transformation of a tech prep biology teacher*. Unpublished doctoral dissertation, The Graduate of a Clemson University, South Carolina, USA.
- Hazır-Bıkmaz, F. (2004). Sınıf öğretmenlerinin fen öğretiminde öz yeterlilik inancı ölçeğinin geçerlik ve güvenilirlik çalışması [Reliability and validity study of a scale to assess classroom teachers' self-efficacy belief toward science teaching]. *Milli Eğitim Dergisi*, 161. Retrieved May 15 2008 <http://yayim.meb.gov.tr/dergiler/161/bikmaz.htm>
- Işıksal M & Aşkar, P (2003). İlköğretim öğrencileri için matematik ve bilgisayar öz yeterlik algısı ölçekleri [The scales of perceived mathematics and computer self-efficacy for elementary students]. *Hacettepe Üniversitesi, Eğitim Fakültesi Dergisi*, 25, 109-118.
- Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 8: Users references guide* (2nd Ed.). Lincolnwood, IL: SSI.
- Lian, M. W. S. (2003). Development of a self-efficacy scale for assessing secondary school students' science self-efficacy beliefs. Retrieved June 5 2008 http://www.ipbl.edu.my/inter/penyelidikan/jurnalpapers/jurnal2003/2003_mary.pdf
- Kelloway, E. K. (1998). *Using LISREL for structural equation modeling: A researcher guide*. Newbury Park, CA: Sage.
- Kupermintz, H. (2002). Affective and conative factors as aptitude resources in high school science achievement. *Educational Assessment*, 8 (2), 123-137.
- Kurbanoglu, S. S. (2004). Öz-Yeterlik inancı ve bilgi profesyonelleri için önemi [Self efficacy belief and its importance for information professionals]. *Bilgi Dünyası*, 5 (2), 137-152.
- Kurbanoglu, S. S., Akkoyunlu, B. & Umay A. (2006). Developing the information literacy self-efficacy scale. *Journal of Documentation*, 62 (6), 730-743.
- Lawson, A. E. Banks, D. L., & Logvin L. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching*, 44 (5), 706-724.
- Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). *SPSS for intermediate statistics; use and interpretation* (Second Edition). Lawrence Erlbaum Associates: Mahwah, NJ.
- Mangos, P. M. & Steele- Johnson, D. (2001). The role of subjective task complexity in goal orientation, self-efficacy, and performance relations. *Human Performance*, 14 (2), 169-186.
- Morgil, İ., Seçken, N. & Yücel, S. S. (2004). Kimya öğretmen adaylarının öz-yeterlik inançlarının bazı değişkenler açısından incelenmesi [Based on some investigation of self-efficacy beliefs of preservice chemistry teachers variables]. *BAÜ Fen Bilimleri Enstitüsü Dergisi*, 6 (1), 62-72.
- Nevill, M. A. (2008). *The impact of reading self-efficacy and the regulation of cognition on the reading achievement of an intermediate elementary sample*. Unpublished doctoral dissertation. Indiana University of Pennsylvania, Indiana.
- Önen, F. & Öztuna, A. (2005). Fen bilgisi ve matematik öğretmenlerinin öz yeterlik duygusunun belirlenmesi [Determining science and mathematics teachers' feeling of self-efficacy]. *Yeditepe Üniversitesi, Eğitim Fakültesi e-Dergisi*. Retrieved August 8 2008 <http://www.istekyasam.com/edu7dergi1/makale2.htm>

- Ören-Şaşmaz, F. & Tatar, N. (September, 2006). Öğretmen adaylarının bazı değişkenler açısından öz-yeterlik inanç düzeylerinin incelenmesi üzerine bir araştırma [A research on exploring prospective teachers' belief of self-efficacy in terms of certain variables]. *Proceedings of the 15th annual conference of the Educational Sciences*: (pp. 403-405). Muğla Üniversitesi, Türkiye.
- Pajares, F. & Kranzler, J. (April, 1995). *Role of self efficacy and general mental ability in mathematical problem solving: a path analysis. Proceedings of the Annual Meeting of the American Educational Research Association*: (pp 1-37). San Francisco.
- Pajares, F. (2000). *Schooling in America: Myths, mixed messages, and good intentions. Lecture delivered at the great teachers lecture series, Cannon Chapel. Emory University, Atlanta.*
- Rose, D.L.W. (2003) *Student self-efficacy in college science: an investigation of gender, age, and academic achievement.* Unpublished Master of Science thesis. University of Wisconsin-Stout, USA.
- Schunk, D.H. (1996, April). *Self efficacy for learning and performance. Presented at American Educational Research Association.* New York.
- Senemoğlu, N. (1997). *Gelişim, öğrenme ve öğretim: Kuramdan uygulamaya.* [Growth, learning and teaching: from theory to practice]. Ankara: Spot Yayıncılık.
- Singer, Y. M. (2008). Physically disabled second graders' writing skills. Retrieved September 18 2008 <http://home.att.net/~ysinger/phws.htm>
- Sörbom, D., & Jöreskog, K. G. (1982). The use of structural equation models in evaluation research. In C. Fornell (Ed.), *Measurement and Evaluation: Vol. 2. A second generation of multivariate analysis* (pp.381-418). New York: Praeger.
- Thompson, R. (2003). *The effect of a laboratory-based, in-context, constructivist teaching approach on preservice teachers' science knowledge and teaching efficacy.* Unpublished doctoral dissertation. University of North Texas, USA.
- Umay, A. (September, 2002). *İlköğretim matematik öğretmenliği programının öğrencilerin matematiğe karşı öz-yeterlik algısına etkisi* [The effect of the primary school mathematics teaching program on the mathematics self-efficacy of students]. *Proceedings of the V. annual conference of the National Science and Mathematics Education.* Orta Doğu Teknik Üniversitesi Eğitim Fakültesi. Ankara. Retrieved September 15 2008 www.fedu.metu.edu.tr/ufbmek5/b_kitabi/PDF/OgretmenYetistirme/Bildiri/t311DA.pdf
- Vural, D. E. & Hamurcu, H. (2008). Preschool teacher candidates' self-efficacy beliefs regarding science teaching lesson and opinions about science. *İlköğretim Online*, 7(2), 456-467.
- Yavuzer, Y. & Koç, M. (2002). Eğitim fakültesi öğrencilerinin öğretmen yetkinlikleri üzerinde bir değerlendirme. An assessment on teacher perfection of education faculty students. *Niğde Üniversitesi, Eğitim Fakültesi Dergisi*, 1 (1), 35-43.
- Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist*, 38 (4), 189-205.

Fen ve Teknolojiye Yönelik Öz-yeterlik Ölçeğinin Geçerlik ve Güvenirlik Çalışması

(Özet)

Problem Durumu: Öz yeterlik bir konuda bireyin düşüncelerinin, davranışlarının, inançlarının ve hissettiklerinin bir ifadesidir. Bireyin karşılaşılabileceği olay ya da olaylar karşısında nasıl bir davranış sergileyebileceğine ilişkin kendi hakkındaki yargısıdır. Öz-yeterlik algısı düşük olan kişiler kendilerine göre zor bir görevle karşılaştıklarında başarılı olmanın yollarını düşünmekten çok karşılarına çıkabilecek engelleri düşünerek çaba göstermekten vazgeçerler. Bu nedenle düşük performans gösterirler. Öz-yeterlik algısı yüksek olanlarsa zor görevler karşısında daha mücadeleci davranırlar ve sorumluluk alırlar. Kişinin öz-yeterlik inancı, onun algısını, tutumunu, motivasyonunu ve performansını farklı şekillerde etkiler. Bireylerin spor, sanat, bilgisayar gibi pek çok alanda öz yeterliklerini inceleyen çalışmalar bulunmaktadır. Bilgi ve teknolojiadaki hızlı gelişmeler eğitim alanında yeni anlayışları doğurmuştur. İçinde bulunduğumuz çağ bilgiyi edinmeyi zorunlu kılmaktadır. Bilgi var olan teknolojiyi daha ileriye götürmekte, teknoloji de bilgiyi aktarma hızını artırmaktadır. Bu sürecin farkına varmış çağdaş toplumlar, yeni nesillerine var olan teknolojiden faydalanarak bilgiye ulaşma yollarını öğretme çabası içerisindeyler. Bu bakımdan okullarda yapılan eğitimin niteliği ve niceliğinde değişimler ortaya çıkmıştır. Ülkemizde de 2004 yılından itibaren eğitim programlarının yapısında köklü değişimler başlatılmıştır. Yenilenen ilköğretim programı incelendiğinde ilköğretim fen bilgisi dersinde içerik, ders saati ve dersin adı başta olmak üzere köklü değişimlerin olduğu dikkati çekmektedir. Ayrıca Fen bilgisi dersine teknoloji boyutunun eklenmiş ve dersin adı "Fen ve Teknoloji" dersi olarak değiştirilmiştir. Ulusal ve uluslararası alan yazında, öğrencilerinin fen bilgisi dersine yönelik, fene yönelik, fen laboratuvarlarına, teknolojiye yönelik öz yeterliklerini belirleyen ölçekler bulunmaktadır. Ancak yenilenen program ile birlikte ilköğretim öğrencilerin fen ve teknolojiye yönelik öz-yeterliklerini belirleyen ölçme araçlarının bulunmadığı görülmektedir. Bu nedenle, fen ve teknoloji eğitiminde öğrencilerin öz-yeterliklerini ortaya koyabilecek bir ölçme aracına gereksinim olduğu düşünülmektedir.

Araştırmanın Amacı: Bu araştırmanın amacı, ilköğretim ikinci kademe öğrencilerinin fen ve teknolojiye yönelik öz-yeterliklerini değerlendirmek amacıyla "Fen ve Teknolojiye Yönelik Öz-yeterlik Ölçeği" (FTÖÖ) geliştirmek ve bu ölçeğin geçerlik ve güvenirlilik çalışmasını yaparak bu ölçeği eğitim alanında öğretmenlerin ve eğitim araştırmacılarının kullanımına sunmaktır.

Araştırmanın Yöntemi: Bu çalışmaya 10 tane ilköğretim okulunun 6. 7. ve 8. sınıflarında öğrenim gören 400 öğrenci katılmıştır. Ölçek hazırlanırken alan yazından faydalanılmış, öğrencilerin ve uzmanların görüşleri alınmıştır. Verilerin analizi için SPSS 10.00 ve LISREL 8.80 paket programları kullanılmıştır. İlk olarak 36 madde içeren FTÖÖ yapı geçerliğini belirlemek

amacıyla açıklayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) uygulanmıştır. Ölçeğin güvenilirliğini belirlemek için madde-toplam korelasyonları, toplam puana göre belirlenmiş üst % 27 ve alt % 27'lik grupların madde puanları arasındaki farkın anlamlılığı için t-testi ve faktörlerin Croanbach alfa iç tutarlılık katsayısı incelenmiştir. Ayrıca, ölçeğin faktör puanlarının ortalama ve standart sapma değerleri incelenmiş, faktörler arasındaki korelasyonlar Pearson Momentler Çarpımı Korelasyon tekniği ile hesaplanmıştır.

Araştırmanın Bulguları: Faktör analizi yapılmadan önce, verilerin faktör analizine uygunluğu Kaiser-Meyer-Olkin (KMO) ve Barlett Sphericity testi ile incelenmiştir. Araştırmada, ölçeğin KMO değeri 0.95 olarak bulunmuştur. Barlett Sphericity test sonucu [6616.83 ($p<0.01$)] olarak bulunmuştur. AFA'ya 36 madde ile başlanmıştır. Son durumda AFA ve DFA analizi sonuçları, 27 madden oluşan FTÖÖ'nin üç faktörlü bir yapıya sahip olduğunu göstermiştir. Bu üç faktör "Fen ve teknolojiye yönelik güven", Fen ve teknoloji ile ilgili zorluklarla başa çıkabilme" ve "Fen ve teknoloji performansına güven" olarak isimlendirilmiştir. Bu faktörler isimlendirilirken alan yazından faydalanılmıştır. Ölçekte yer alan faktörlerin iç tutarlılık katsayıları sırasıyla, 0.93, 0.75 ve 0.80 olarak bulunmuştur. Ölçeğin bütününe ilişkin Croanbach alpha katsayısı 0.93'tür.

Araştırmanın Sonuçları ve Önerileri: Bu çalışmada araştırmacılar tarafından geliştirilen ve geçerlik güvenilirlik çalışması yapılan "Fen ve Teknolojiye Yönelik Öz yeterlik" ölçeği tanıtılmıştır. Hazırlanan ölçek ilköğretim öğrencilerinin fen ve teknolojiye yönelik öz-yeterliklerini belirlemek için kullanılmaya uygun bir ölçektir. Alandaki çalışmalar incelendiğinde, fen ve teknolojiye yönelik öz yeterlik ölçeğinin az sayıda geliştirildiği görülmektedir. Geçerlik ve güvenilirlik çalışması yapılan ölçeğin bu alanda yapılan araştırmalara önemli katkılar sağlayacağı düşünülmektedir. Geliştirilen ölçek ilköğretim öğretmenlerin öğrencilerinin fen ve teknoloji dersine yönelik öz-yeterliğini belirlemede kullanılabilir. Öz yeterlik inancının tutum, algı, motivasyon gibi pek çok değişkeni etkilediği düşünüldüğünde öğretmenlerin derslerinde bu ölçeği kullanmalarının gerekliliği bir kez daha ortaya çıkmaktadır. İleriki araştırmalarda ise FTÖÖ ölçeği kullanılarak farklı değişkenlerle yapılacak korelasyonel çalışmalarda (örneğin, akademik başarı, cinsiyet, tutum, vb.) öğrencilerin öz yeterlik ilişkileri irdelenebilir.

Anahtar Sözcükler: İlköğretim öğrencileri, fen ve teknoloji eğitimi, öz-yeterlik, ölçek geliştirme

Appendix A

FEN VE TEKNOLOJİ DERSİ ÖZYETERLİK ÖLÇEĞİ

Açıklama: Bu ölçekte, Fen ve teknoloji dersine ilişkin düşüncelerinizi belirlemek için her cümle için karşısında TAMAMEN KATILYORUM, KATILYORUM, KARARSIZIM, KATILMIYORUM ve HİÇ KATILMIYORUM olmak üzere beş seçenek verilmiştir. Her cümleyi dikkatle okuduktan sonra kendinize en uygun gelen seçeneği işaretleyiniz.

	Tamamen Katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Kesinlikle Katılmıyorum
Fen ve teknoloji dersindeki problemler beni endişelendirir.					
Fen ve teknoloji problemlerini çözerken zorlanırım.					
Fen ve teknoloji sınavları beni endişelendirir.					
Fen ve teknoloji dersinde araştırma ödevi almak istemem.					
Fen ve teknoloji ödevlerimi tek başıma yapamam.					
Ne kadar çaba harcasam da fen ve teknolojiyi öğrenemem.					
Fen ve teknoloji konularını anlamakta zorlanan arkadaşlarıma yardım edebilirim.					
Fen ve teknoloji öğretmenim sorduğu soruları cevaplayamamaktan korkarım.					
Fen ve teknoloji deneylerinde sonuca ulaşamamaktan her zaman korkarım.					
Fen ve teknoloji dersinde zorlandığımda bu zorluğun üstesinden tek başıma gelebilirim.					
Fen ve teknoloji dersinde başarılı olmak için gerekli becerilere sahibim.					
Eğer seçim hakkım olsaydı, fen ve teknoloji dersini öğrenmek istemezdim.					
Fen ve teknoloji projelerini başarı ile tamamlayabilirim.					
Fen ve teknoloji konuları ister zor, ister kolay olsun, bu konuları anlayabileceğimden eminim.					
Zor olan fen ve teknoloji kavramlarını anlayabileceğimden çok emin değilim.					
Fen ve teknoloji sınavlarında başarılı olacağımdan eminim.					
Ne kadar çabalarsam çabalayayım, fen ve teknoloji konularını öğrenemiyorum.					
Fen ve teknoloji ile ilgili etkinlikler çok zor olduğunda, bunları yapmaktan vazgeçerim veya sadece kolay kısımlarını yaparım.					
Fen ve Teknoloji Dersinden yüksek not alacağıma inanıyorum.					
Fen ve Teknoloji Dersinde anlatılan temel kavramları anlayabileceğim konusunda kendime güveniyorum.					
Fen ve Teknoloji Dersinde öğretmenin anlatacağı en zor konuyu bile anlayacağıma inanıyorum.					
Fen ve Teknoloji Dersindeki ödevleri ve sınavları mükemmel yapabileceğim konusunda kendime güveniyorum.					
Fen ve Teknoloji Dersinde başarılı olmayı bekliyorum.					
Eminim ki Fen ve Teknoloji Dersinde öğretilen tüm becerileri ustalıkla yapabilirim.					
Fen ve Teknoloji konularında verilen görevleri tamamlayabilirim.					
Fen ve Teknoloji konularında kendime güvenerek çalışırım.					
Fen ve Teknoloji konularında kendimi geliştirebilirim.					

SELF EFFICACY SCALE TOWARDS SCIENCE AND TECHNOLOGY

	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Problems relevant to science and technology lesson disconcert me.					
I try hard to solve science and technology problems.					
Science and technology tests disconcert me.					
I hesitate to take homework requiring research in science and technology lessons.					
I cannot make my science and technology homework by myself.					
I cannot learn matters of science and technology no matter how I spend effort.					
I can help my friends, who have difficulties in understanding science and technology matters.					
I fear not to be able to answer questions of my teacher about science and technology.					
I always fear not to be able to achieve to the conclusion in science and technology experiments.					
I can cope with difficulties by myself in science and technology lesson whenever I encounter.					
I have skills required for being successful in science and technology lesson.					
If it was optional, I would not want to learn science and technology lesson.					
I can accomplish science and technology projects successfully.					
I am sure that, I can understand science and technology matters whether they are easy or not.					
I am not very sure that, I can understand difficult scientific and technologic concepts.					
I am sure that, I will be successful in science and technology examinations.					
I cannot learn science and technology matters no matter how I try hard.					
If scientific and technologic activities are hard, I give them up or I accomplish parts, which are easier.					
I believe that, I will have high scores in science and technology examinations.					
I am self-confident about that, I can understand basic concepts mentioned in science and technology lessons.					
I believe that, I will understand even the most difficult matter, which will be taught by science and technology teacher.					
I am self-confident about that; I can accomplish homework and examinations in science and technology perfectly.					
I am expecting that, I will be successful in science and technology lesson.					
I am sure that, I can accomplish all skills given in science and technology lesson successfully.					
I can accomplish the tasks given in science and technology matters.					
I study on science and technology matters self-confidently.					
I can develop myself on science and technology matters.					