The Effect of Using Alternative Assessment Activities on Students' Success And Attitudes in Science and Technology Course

Esma BULUŞ KIRIKKAYA[a]
Kocaeli University

Gurbet VURKAYA
Science and Technology Teacher

Abstract

The pre-test and post-test quasi-experimental design with control group was used in this study, in which the impact of alternative assessment activities on students' academic achievement levels and attitudes were explored by employing these activities in the unit "Electricity in Our Lives" of the Science and Technology Course. The research was carried out in three elementary schools in the city of Kocaeli, which exhibited three different levels of success in the SBS (National Level Determination Exam). The academic achievement test, which was one of the data collection tools of the study, was developed by the researchers and its reliability was found to be .85. A Likert-type scale, whose reliability was found to be .96, was employed in order to measure students' attitudes towards science. In addition, 14 alternative assessment activities were used in the study, which were developed by the researchers and composed of structural communication grid, diagnostic tree and predict-observe-explain activities. According to the findings of the research, a significant difference was found between the experimental and the control groups with respect to science attitudes and achievement in the favor of the experimental groups. A generally positive correlation was found between the mean scores that the students in the experimental group got from the activities and their achievement pre-test-post-test scores.

Key Words

Science and Technology Course, alternative Assessment Activities, Structural Communication Grid, Diagnostic Tree, Predict-Observe-Explain.

Osborne and Wittrock (1983) underlined that the knowledge a student or an individual possesses is of importance for responding new information. This opinion was based to constructivist learning approach. Brooks and Brooks (1999) suggested that the student works actively in this process to construct knowledge and that she/he has self-control (cited in Hançer, 2006). In an educational environment in which students actively seek to construct knowledge, it is not enough to use only the conventional assessment approaches. Students either interpret the new situation or prefer to undertake regulations again in the event that structuring is not meaningful. (Brooks & Brooks, 1993).

Colburn (2000), effectively using the principles of the constructivist approach, makes some suggestions for the teaching of science and technology. First one is the inquiry method of the central of teaching science, second one, is to encourage collaborative learning in class, and the last one is importance of questions asked in classroom.

Alternative assessment approaches are used to assess students' knowledge and skills, and they are different from conventional methods. Alternative assessment is a philosophy and it is goal-oriented. Alternative assessment generally emerged out of a practice connected to an educational reform and of a new quest (Donovan, Larson, Stechschulte, & Taft 2002; McMillan 2001; Settlage, 2004). Stone (1990)

argued that new practices and strategies are needed for the selection of a good assessment method, and underlined the importance of using alternative assessment methods instead of conventional ones.

In their studies, Herman, Aschbacher and Winters (1992) listed the common features of alternative evaluation as enabling students to do something new, to develop their higher thinking skill on the basis of comprehension, to solve problems, to take responsibility and to undertake tasks, to have a real world interaction and to make more coherent decisions, i.e. to make a better evaluation (cited in Corcoran, Dershimer, & Tichenor, 2004). In their studies, Conradie and Frith (2000) and Bol, Ross, Nunnery and Alberg (2002) maintained that the use of alternative methods would make evaluation flexible and yield positive results.

The structural communication grid (SCG) technique, which is among alternative assessment activities, is an assessment technique aiming at revealing students' cognitive structures and is composed of small boxes. In this assessment technique that is different from the multiple choice test technique, students are expected to define the boxes that convey the most suitable responses and to number the boxes in a correct order. Therefore, this technique is called "structural communication grid" (Jonstone, Bahar, & Hansell, 2000). A structural grid consists of rows and columns and one can place pictures, figures, scripts etc. in the bricks depending on the level of students. Furthermore, the number of bricks might vary depending on the group (Durmuş & Karakırık, 2005). This flexibility is thought to enable the person using structural communication grid to make a more objective assessment and evaluation over a wider population (Bahar 2003, Durmuş & Karakırık, 2005). This technique is of great importance since it measures meaningful learning and determines the misconceptions in the information network (Jonstone et al., 2000). Solas (1992) underscores the importance of using the structural communication grid for assessment purposes. Durmuş and Karakırık (2005) argued that it is an alternative to multiple-choice tests.

In diagnostic tree, students are asked to make the best choice by choosing among true or false statements in a way that is from the most basic to the most detailed one. In this way, one obtains a tree with 8 or 16-choice list of expressions (MEB, 2006). The technique is considered an alternative to the traditional true-false tests. Its advantages include determining fallacies, enabling on-computer prep-

aration and providing an opportunity to notice and correct any wrong decisions. Nevertheless, the most significant limitation of the technique is that it proves to be incapable of assessing higher level learning skills (Bahar, Nartgun, Durmuş, & Bicak, 2008).

The predict-observe-explain (POE) method was developed by White and Gunstone (1992). This method consists of three phases and it is implemented by taking the opinions of each student and developing reasons. It is also referred to in the literature as POE (Prediction- Observation- Explanation) method. There exist studies in which this method was employed to investigate children's understanding of science concepts (White & Gunstone, 1992; Liew & Treagust, 1995). In the diagnostic tree (DT) method, students are asked to pick correct answers among expressions aligned from the most general to the most detailed. Thus, a branched tree culminating in a list of expressions with eight or sixteen choices emerges (National Ministry of Education [MEB], 2006). This technique is seen as an alternative to conventional true/false tests. Among the advantages of this method are determinability of incorrect knowledge, production at a computer environment and the possibility of detecting and correcting wrong decisions.

Gamor (2001) revealed in her study that alternative assessment and evaluation methods have facilitated the construction of knowledge and phenomena compared to the conventional assessment methods. In the study carried out by Andrade and Du (2005), the importance of alternative assessment practices for student achievement is underlined. Janisch, Liu and Akrofi (2007) indicated in their study that alternative assessment methods will help students acquire awareness about their needs, follow academic developments, and understand the values regarding the learning environment and the learning process. Köseoğlu, Tümay and Kavak (2002) observed, at the end of their study in which they employed the predict-observe-explain method, that students' attitudes towards chemistry improved and their motivation levels went up.

In his study, Karahan (2007) used "structural grids", "concept maps" and "diagnostic trees" and made an attempt to determine the effect of alternative assessment and evaluation techniques on achievement. He concluded that alternative assessment and evaluation techniques had a positive influence on students' achievement level. Kartal and Buldur (2007) conducted a study with the aim of determining in which alternative evaluation technique(s) students

were more successful and compared their scores in alternative evaluation techniques with those in traditional assessment and evaluation methods. Their study concluded that students were most successful and least successful in "diagnostic tree" and "structural grid" respectively. Moreover, they were more successful in exams devised through alternative evaluation techniques than traditional ones. Nevertheless, Orhan (2007) found that teachers used diagnostic tree less often than any other alternative evaluation methods. Similarly, Ersoy (2008) discovered that the least commonly used technique was structural communication grid. Karacak-Deren (2008) observed a significant variation in the level of achievement obtained by students from an achievement test consisting of structural grids, matching, puzzle.

Similarly, worksheets were found to have a positive impact on attitude (Bozdoğan, 2007). İnce (2007) investigated the effect of portfolios on 6th grade students' attitudes towards science and technology and on reducing their text anxiety. The researcher concluded that portfolios had a positive effect on students' attitudes. Mıhladız (2007) also found that portfolios affected students' attitudes in a positive way.

In this study, the aim is to explore the impact of the use of structural communication grid, diagnostic tree and predict-observe-explain methods on student achievement and attitude. When the studies on the subject were reviewed, no study was found on 7th grade students in terms both of the units selected and of the methods employed. Therefore, it is believed that this study will be compelling with respect to its findings.

Answers were sought to the following sub-problems by departing from the problem: "Does the use of alternative assessment activities in the unit "Electricity in Our Lives" have any significant impact on students' academic achievements and attitudes?

1. Is there any significant difference between the experimental and the control groups in terms of the scores of students in science attitude and achievement post-tests?

2. Is there any significant difference between the science attitude and achievement pre-test and post-test scores of the experimental and the control group students?

3. Is there any significant difference between the achievement pre-test and post-test scores of experimental and control group students with respect to their school achievement levels?

4. Is there any significant difference between the scores that the experimental group students got from the activities prepared by using structural communication grid, diagnostic tree and predict-observe-explain techniques, with respect to school success level?

5. Is there any significant correlation between students' achievement test scores and the scores they got from the activities prepared by using structural communication grid, diagnostic tree and predict-observe-explain techniques?

Method

Model

Quasi-experimental method was used in this study in which the impact of alternative assessment activities on students' attitudes and academic achievement were explored. The difference of the quasi-experimental method from the true-experimental method is the selection of groups whose levels are appropriate to each other before the selection (Karasar, 2006). The pre-test, post-test quasi-experimental design with control group was used in this study.

Universe and Sample

The universe of the research consists of all 7th grade students in the central district of Kocaeli. To determine the sample of the research, the elementary schools located in the area were ranked by researchers according to their success levels in the SBS 2007 exam, and then they were divided into three strata each of which consisted of equal number of schools. Of the schools in each stratum, the average SBS achievement was calculated and two schools in each stratum that are the closest to those averages were defined. Then, six schools were selected whose success levels were low, medium and high. With the evaluations performed after the level-determination and science attitude scales were administered, two classrooms whose success levels and science attitudes were similar to each other were included in the sample in one of the two schools placed in each stratum. In this study, the stratification technique was employed at the stage of determining schools, and the random sampling technique was used at the stage of defining the sample where the study would be conducted in the school. The study group consists of 132 students (66 female, 66 male) enrolled in these schools in the 2008-2009 Academic Year.

When the level-determination, attitude and achievement pre-test results of the students enrolled in the schools of A, B and C were examined, no significant difference was found between the experimental and control groups (*p>.05*). It could then be concluded that the experimental and control groups in each school were similar to each other in terms of prior-knowledge levels and attitudes.

Level-Determination Test: The assessment questions in the unit "Electricity in Our Lives" of the 6th grade Science and Technology textbook and workbook, published by the Ministry of Education, were rearranged without disrupting their authenticity and used for level-determination purposes. The appropriateness of the questions to the acquirements and their suitability to content validity were given special attention. In this test, a total of 25 questions were included: 10 true/false questions, 5 gap-filling questions, and 10 multiple choice questions.

Science Attitude Test: The scale developed by Demirci (2003), which consists of 32 items on a 5-point Likert-type scale was used.

The scale's Cronbach's alpha reliability coefficient was found by Demirci to be .96. Özçelik (2007) used this scale and calculated its alpha value to be .93. The scale was found to be suitable for this study since the reliability values obtained were high. Whereas α value of the scale was found in pre-test in this study to be .91, it was .94 in the post-test..

Achievement Test: The academic achievement test was developed by the researchers in order to determine students' academic achievement levels in the unit "Electricity in Our Lives". During the test development process, a pool was first formed, which consisted of 90 multiple-choice questions that had been prepared in line with the acquirements. According to the opinions of experts, 50 questions that comprised the acquirements were selected out of this pool. A pilot study was first carried out in order to determine the validity and reliability of the achievement test. The pilot study was carried out with 300 8th grade students who were enrolled in four elementary schools, which were selected among those schools that had not been included in the research in the city of Kocaeli. The aim of administering the achievement test on 8th grade students is that these students already possess these acquirements. For the item analysis of the test, all scores were ranked from the highest to the lowest. The best 27% of all the papers were put into the "higher group" and the worst 27% were put into the "lower group", and the rest of the papers were not considered. The most difficult and the

questions were excluded and the number of questions was reduced to 25, whose content validity was tested. Of the achievement test that consisted of 25 questions, the mean item difficulty was found to be .50, and the discriminating power of the test was found to be .56. In the study, both the split-half method and the KR method were employed as reliability measurement techniques. The Pearson's correlation between the odd and even numbered items was r=.77. This is a positive significant correlation, that is, the test's internal consistency is high (Aydoğdu & Kesercioğlu, 2005). Pearson's correlation values between the first and the last half of the test were found to be r=.70. This is also a positive significant correlation. Finally, the test's reliability of measurement was tested according to KR-20 and it was found to be r_x=.85. Measurement reliability of achievement test is greater than r_x=.70, that means that the reliability is high (Yılmaz, 1997).

Alternative Assessment Activities: The study involved 14 different alternative assessment activities. For the validity and reliability of the activities developed by the researchers, the designed alternative assessment activities (SCG, DT and POE) were shared with fourth-grade Science Teaching students at Kocaeli University for six class hours in two different days and, thus, the activities' content scope, difficulty level and responses to be given to questions were discussed. The activities were rearranged according to the outcomes of these discussions. After these arrangements, the opinions of three field experts and six Science and Technology teachers were requested in order to ensure the activities' validity and reliability. The activities were finalized in line with the opinions after this study that approximately lasted six weeks. Besides, in line with the experts' opinions and by also considering the activities' durations of implementation, the number of activities was reduced.

Results and Discussion

The findings of this study obtained by exploring the impact of the use of alternative assessment activities in the unit "Electricity in Our Lives" on students' attitudes and academic achievement levels were discussed with respect firstly to its impact on students' attitudes and then to its impact on their achievements. The discussion, which continued with the success of the control groups selected from schools at different success levels in alternative assessment activities, ended with the comparison of the success demonstrated in alternative assessment activities within same groups with the achievement test results.

According to the conclusions drawn from the findings, a difference was found between the experimental and control groups in each of the schools A, B and C in terms of the scores obtained from the science attitude post-test. Whereas the difference between the experimental and the control group in the school, A was not significant, the differences in other schools were significant (A: $p>.05$, B: $p<.01$, C: $p<.01$). The finding that the differences between the post-test scores in the schools with medium and low success levels were significant whereas the differences between the experimental and the control groups in each school in terms of the science attitude pre-test scores were insignificant might indicate that the assessment activities carried out positively influenced the science attitude in these schools. When the studies that explore the impact of different methods employed in classes on students' attitudes were reviewed, instances were found where the attitude post-test scores of experimental and control groups differed significantly. Studies carried out by Gazioğlu (2006), Ceylan (2008), Yener (2006) and Güçlüer (2006) can be given as examples.

Significant differences were found between the science attitude pre-test and post-test scores in the control groups in all schools and in the total of experimental groups (Experimental Group 1: $p<.01$, Experimental Group 2: $p<.01$, Experimental Group 3: $p<.01$ and Experimental Total Group $p<.01$). Of the students placed in experimental groups, the mean scores obtained from the science attitude post-test are significantly higher than the mean scores obtained from the science attitude pre-test. Therefore, it could be argued that the alternative assessment activities carried out on experimental group students took effect and elicited a positive change in their attitudes towards science. Studies carried out by Hamayan (1995), Conradie and Frith (2000), Janisch et al. (2007), which indicated that alternative assessment activities might generate a positive improvement in students' attitudes towards the course taught, point to similar findings.

Differences were found between the science attitude pre-test and post-test mean scores of the total control groups in the schools A, B and C. However, it is observed that only the difference between the pre-test and post-test scores of the control group in the school with high success level is significant. This finding might give rise to the idea that achievement positively influenced attitudes independent from the methods employed. In a research that

investigated the relationship between academic achievement and attitudes, it was found that there exists a positive medium-level correlation between the attitudes of science teacher candidates towards the chemistry course and their academic achievement levels (Hançer, Uludağ, & Yılmaz, 2007). No significant difference was found between the pretest and post-test scores in the control groups in schools with medium and low success levels. In a study conducted by Dinçer, Ece and Yıldızlar (2008), it is suggested that there exists a difference between the attitude pre-test scores and post-test scores of the control group, but that this difference is not significant although a clear increase is observable in attitude scores.

A significant difference was found between the experimental and control groups in the schools A, B and C with respect to the scores obtained from the achievement post-test (A: $p<.01$, B: $p<.01$; C: $p<.05$). It was observed that the students of the experimental group scored higher than the students of the control group. This finding might be interpreted in a way that the alternative assessment activities that were performed had a positive impact on students' achievement. When the studies that explored the impacts of different methods implemented alongside conventional methods on student achievement are reviewed; while no difference is observed in students' science achievements in the studies carried out by Ceylan (2008), Ünal and Ergin (2006); studies conducted by Yılmaz and Huyugüzel-Çavaş (2006), Demirbaş and Yağbasan (2006), Yıldız-Duban (2006) Kıncal, Ergül and Timur (2007) concluded that the different methods and techniques, when compared to conventional ones, had positive impacts on achievement.

When the t-test results related to the achievement pre-test and post-test scores of the experimental and control groups were examined; significant differences were observed between the pre-test and post-test total scores of both the experimental group in each school and the total experimental group (Experimental Group 1: $p<.01$, Experimental Group 2: $p<.01$, Experimental Group 3: $p<.01$ and Experimental Total Group: $p<.01$). It could therefore be said that the scores that the control groups got from the achievement post-test are significantly higher than their scores in the achievement pre-test. This finding might give rise to the idea that the alternative assessment activities positively influenced students' achievement levels. Liew (2004), in the study that investigated the effectiveness of POE activities, attempted to determine

the impact of POE activities on students' science conceptions and achievement levels. As a result, it was observed that even those students exhibiting negative attitudes expressed their opinions using the predict-observe-explain method. The study suggests that the method is influential in determining student achievement.

There exist differences between the pre-test and post-test scores of all the three control groups. Whereas the differences were not significant between achievement pre-test and post-test total scores of the Control 1 group in the school A and the Control 3 group in the school C (Control Group 1: *p>.05,* Control Group 3: *p>.05*), the difference was significant between the pre-test and post-test scores of the Control 2 group in the school B (Control Group 2: *p<.01*). Increase in students' achievement levels is expectable since the contributions of teaching activities no matter how discordant the process is are non-negligible. Numerous studies reported increases in students' achievement levels that were not statistically significant. Similar results were observed in the studies carried out by Kalem, Tanel and Çallıca (2002) and Balım, İnel and Evrekli (2008). In some other studies, on the other hand, it was suggested that achievement post-test scores of control group students were found to be higher than pre-test scores and these differences were significant. Studies of Taşdemir and Tay (2007) and Orhan and Bozkurt (2009) support this result.

When the in-school variance of the achievement pre-test scores of experimental and control groups was examined, a significant difference was observed (*p<.01*) and it was determined that the school A, the success level of which was high, was significantly more successful than the schools B and C. It was also determined that the achievement post-test scores of experimental groups ranged according to the success levels of schools, and that the difference between the school groups was significant, similar to that in the pre-test, in the favor of the school with high success level (*p<.01*). In the control groups where the achievement post-test was performed, it was observed that the results ranged according to the success levels of schools but that no significant difference existed (*p>.05*).

When the scores that the control groups got from the activities prepared using structural communication grid, diagnostic tree and predict-observe-explain techniques were examined in terms of schools' success levels; it was found that there exists a significant difference. In order to determine the

degree and location of this difference, Tukey test analysis was performed to examine the mean scores obtained from SCG, DT and POE techniques, and it was found that the school B scored significantly higher than the schools A and C (*p<.05, p<.01*). Similarly, the school A scored significantly higher than the school C (*p<.01*). When the total scores obtained from the SCG, DT and POE techniques were analyzed with respect to schools' success levels, it was observed that the school success level created significant difference, and that the control group students in the school B scored the highest in terms of activity averages. Although it was tried in this research to select practice teachers in experimental groups from the same level and the necessary instructions and warnings were delivered both before and after the performance of the activities; it was not possible to level teachers' attitudes towards students. The teacher at the school B became a significant factor in the easy and successful execution of the student communication activities. In order to eliminate this effect, it could be suggested for future studies that the researcher himself/herself should conduct the class in experimental and control groups. However, another point of concern is the difficulty to receive permission from both administrators and teachers for the researcher to conduct the classes for such long times.

When the relationship between students' scores in the alternative assessment activities and their scores in the achievement test was analyzed through the Pearson's correlation, a positive high-level correlation was found between the mean scores obtained from the SCG activities and the mean scores obtained from the achievement post-test at the level of r= .53. A positive low-level correlation was determined between the mean scores in DT activities and the achievement post-test scores at the level of r= .19. A positive high-level correlation was found between the mean scores in POE activities and the achievement post-test scores at the level of r= .44. A generally positive correlation was determined between the mean scores obtained from the activities and the achievement post-test. Although the size of the sample does not allow for generalization since this was an experimental study, it could still be concluded that those students who successfully perform the assessment activities like structural communication grid and predict-observe-explain show success in the academic achievement test too. In a study conducted by Barootchi and Keshavarz (2002), a correlation was examined between the scores obtained through portfolio, which is one of the alternative assessment techniques, and

the scores obtained from the exams prepared by teachers. The findings of this research indicated that students responded positively to the portfolio technique and that it positively contributed to their achievements in their courses.

Suggestions

Although the fact that the findings cannot be generalized since this study was a quasi-experimental research limits the suggestions; to guide future studies, what follows are several suggestions for researchers and authorities.

1. Researchers could be suggested to carry out studies in which the impacts of several variables like the access to facilities, socioeconomic status and schools' environments (village, small town, city etc.).

2. As seen in studies conducted, different activities performed positively influence students' attitudes towards courses. Comparisons could be made by conducting different teaching and assessment-evaluation methods especially in classes with medium and low success levels.

3. By considering the positive role of the use of different assessment methods and techniques in increasing students' achievement levels, teachers could be encouraged to employ these methods and techniques.

References/Kaynakça

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Assessment, 10* (3), 1-11.

Aydoğdu, M. ve Kesercioğlu, T. (2005). *İlköğretimde fen ve teknoloji öğretimi.* Ankara: Anı Yayıncılık.

Bahar, M. (2003). A study of pupils'ideas about the concept of life. *Kastamonu Eğitim Dergisi, 11* (1), 93-104. Bahar, M., Nartgün, Z., Durmuş, S. ve Bıçak, B. (2008). *Geleneksel alternatif ölçme ve değerlendirme: Öğretmen el kitabı* (2. bs.). Ankara: Pegem A.

Balım, A. G., İnel, D. ve Evrekli, E. (2008). Fen öğretiminde kavram karikatürü kullanımının öğrencilerin akademik başarılarına ve sorgulayıcı öğrenme becerileri algılarına etkisi. *İlköğretim Online, 7* (1), 188-202.

Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests. *Educational Research, 44* (3), 279-288.

Bol, L., Ross, S. M., Nunnery, J. A., & Alberg, M. S. (2002). A comparison of teachers'assessment practices in school restructuring models by year of implementation. *Journal of Education for Students Placed at Risk, Lawrence Erlbaum Associates, Inc, 7* (4), 407-423.

Bozdoğan, A. (2007). *Fen bilgisi öğretiminde çalışma yaprakları ile öğretimin öğrencilerin fen bilgisi tutumuna ve mantıksal düşünme becerilerine etkisi.* Yayınlanmamış yüksek lisans tezi, Çukurova Üniversitesi Sosyal Bilimler Enstitüsü, Adana.

Brooks, J. G., & Brooks, M. G. (1993). *In search for understanding: The case for constructivist classrooms.* Alexandria, VA: Association for Pervision and Curriculum Development.

Ceylan, H. (2008). *İlköğretim fen ve teknoloji dersinde altıncı sınıf öğrencilerine elektrik konusunun öğretiminde kavramsal değişim yaklaşımının öğrenci başarısı ve tutumuna etkisi.* Yayınlanmamış yüksek lisans tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Colburn, A. (2000). An inquiry primer. *Science Scope, 23* (6), 42-45.

Conradie, J., & Frith, J. (2000).Comprehension tests in mathematic. *Educational Studies in Mathematics, Published by: Springer Stable, 42* (3), 225-235.

Corcoran, C. A. Dershimer, E. L., & Tichenor, M. S. (2004). A teacher's guide to alternative assessment taking the first steps. *The Clearing House, 77* (5), 213-216.

Demirbaş, M. ve Yağbasan, R. (2006). Fen bilgisi öğretiminde sosyal öğrenme teorisine dayalı öğretim etkinliklerin, öğrencilerin akademik başarılarına olan etkisinin incelenmesi. *Kastamonu Eğitim Dergisi, 14* (1), 113-128.

Demirci, C. (2003). *Fen bilgisi öğretiminde etkin öğrenme yaklaşımının erişi, tutum ve kalıcılığa etkisi.* Yayınlanmamış doktora tezi, Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Dinçer, M., Ece, A. S. ve Yıldızlar, M. (2008). *İlköğretim okullarında müziklendirilmiş matematik oyunlarıyla yapılan öğretimin akademik başarıya ve tutuma etkisi.* http://oc.eab.org.tr/egtconf/pdfkitap/pdf/434.pdf adresinden 14 Aralık 2009 tarihinde edinilmiştir.

Donovan, R., Larson, B., Stechschulte, D., & Taft, M. (2002). *Building quality assessment.* Washington, DC: Education Resources Information Center (ERIC Document Reproduction Service No. ED466690)

Durmuş, S., & Karakırık, E. (2005). A computer assessment tool for structural communication grid. *The Turkish Online Journal of Educational Technology, 4* (4), 3-7.

Ersoy, E. (2008). *İlköğretim I. kademe fen ve teknoloji dersindeki ölçme ve değerlendirme uygulamasının değerlendirilmesi.* Yayınlanmamış yüksek lisans tezi, Çanakkale Onsekiz Mart Üniversitesi Sosyal Bilimler Enstitüsü, Çanakkale.

Gamor, K. I. (2001). *Moving virtuality into reality: A comparison study of the effectiveness of traditional and alternative assessments of learning in a multisensory, fully immersive physics program.* Doctoral Dissertation, George Mason University, USA.

Gazioğlu, G. (2006). *İlköğretim 7.sınıf öğrencilerinin basınç konusunu kavramada çoklu zeka tabanlı öğretiminin öğrenci başarısı, tutum ve öğrenilen bilgilerin kalıcılığına etkisi.* Yayınlanmamış yüksek lisans tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Güçlüer, E. (2006). *İlköğretim fen bilgisi eğitiminde kavram haritaları ile verilen bilişsel desteğin başarıya hatırda tutmaya ve fen bilgisi dersine ilişkin tutuma etkisi.* Yayınlanmamış yüksek lisans tezi, Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.

Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Rewiew of Applied Linguistics,15,* 212-226.

Hançer, A. H. (2006) Enhancing learning though constructivist approach in science education. *International Journal of Environmental and Science Education, 1* (2), 181-188.

Hançer, H. A., Uludağ, N. ve Yılmaz, A. (2007). Fen bilgisi öğretmen adaylarının kimya dersine yönelik tutumlarının çeşitli değişkenlere göre değerlendirilmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 32,* 100-109.

İnce, E. (2007). *İlköğretim öğrencilerinin fen ve teknoloji dersine karşı olumlu tutum geliştirmelerinde ve sınav kaygısının giderilmesinde portfolyo tekniğinin etkisi.* Yayınlanmamış yüksek lisans tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Janisch, C., Liu, X., & Akrofi, A. (2007). Implementing alternative assessment: opportunities and obstacles. *The Educational Forum, 71* (3), 221-230.

Johnstone, A. H., Bahar, M., & Hansell, M. H. (2000). Structural communication grids: A valuable assessment and diagnostic tool for science teachers. *Journal of Biological Education, 34* (2), 87-89.

Kalem, R., Tanel, Z. ve Çallıca, H. (2002, Eylül). *Ortaöğretim fizik dersi ısı ve sıcaklık konusu öğretim programı geliştirme üzerine bir çalışma.* V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi'nde sunulan bildiri, ODTÜ, Ankara.

Karacak-Deren, Ş. (2008). *İlköğretim 8. Sınıf genetik ünitesinin 5E modeline göre tasarlanan multi-medya destekli öğretiminin öğrencilerin erişi ve tutumuna etkisi.* Yayınlanmamış yüksek lisans tezi, Muğla Üniversitesi, Fen Bilimleri Enstitüsü, Muğla.

Karahan, U. (2007). *Alternatif ölçme ve değerlendirme metodlarından grid, tanılayıcı dallanmış ağaç ve kavram haritalarının biyoloji öğretiminde uygulanması.* Yayınlanmamış yüksek lisans tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Karasar, N. (2006). *Bilimsel araştırma yöntemi kavramlar, ilkeler, teknikler* (16. bs.). Ankara: Nobel.

Kartal, S. ve Buldur, S. (2007, Eylül). *Fen öğretiminde alternatif değerlendirme tekniklerinin kullanılması durumunda öğrencilerin başarı düzeylerinin incelenmesi.* 16. Ulusal Eğitim Bilimleri Kongresi'nde sunulan bildiri, Gaziosmanpaşa Üniversitesi Eğitim Fakültesi, Tokat.

Kıncal, R. Y., Ergül, R. ve Timur, S. (2007). Fen bilgisi öğretiminde işbirlikli öğrenme yönteminin öğrenci başarısına etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 32,* 156-163.

Köseoğlu, F., Tümay, H. ve Kavak, N. (2002, Eylül). *Yapılandırıcı öğrenme teorisine dayanan etkili bir öğretim yöntemi tahmin et- gözle-açıkla "buz ile su kaynatılabilir mi?* V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi'nde sunulan bildiri, ODTÜ, Ankara.

Liew, C., & Treagust, D. F. (1995). A Predict-observe-explain teaching sequence for learning about students' understanding of heat and Expansion of liquids. *Australian Science Teachers' Journal, 41,* 68-71.

Liew, C. W. (2004) *The effectiveness of predict-observe-explain technique in diagnosing students' understanding of science and identifying their level of achievement.* Unpublished doctoral dissertation, Curtin University of Technology, Australia.

McMillan, J. H. (2001). *Essential assessment conceptsfor teachers and administrators.* Thousand Oaks, CA: Corwin Press, Inc.

Mıhladız, G. (2007). *İlköğretim fen bilgisi öğretiminde portfolyo uygulamasının öğrencilerin akademik başarılarına ve derse yönelik tutumlarına etkisi.* Yayınlanmamış yüksek lisans tezi, Muğla Üniversitesi, Fen Bilimleri Enstitüsü, Muğla.

Milli Eğitim Bakanlığı [MEB]. (2006). *İlköğretim Fen ve Teknoloji Dersi (6,7,8. sınıflar için) Öğretim Programı.* Ankara: MEB yayıncılık.

Orhan, A. T. (2007). *Fen eğitiminde alternatif ölçme ve değerlendirme yöntemlerinin ilköğretim öğretmen adayı, öğretmen ve öğrenci boyutu dikkate alınarak incelenmesi.* Yayınlanmamış doktora tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Orhan, A. T. ve Bozkurt, O. (2009). Yapılandırmacı yaklaşıma göre fotosentez konusunun öğretiminin incelenmesi. *Kastamonu Eğitim Dergisi, 17* (3), 905-918.

Osborne, R., & Wittrock, M. C. (1983). Learning science: A generative Process. *Science Education, 67* (4), 489-508.

Özçelik, A. (2007). *İşbirliğine dayalı öğrenmenin, fen bilgisi dersinde başarı tutum ve kalıcılığa etkisi.* Yayınlanmamış yüksek lisans tezi, Eskişehir Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Settlage, J. (2004). Preparing new science teachers for urban classrooms: Consensus within an expert community. *School Science and Mathematics, 104* (5), 214-225.

Solas, J. (1992). Investigating teacher and student thinking about the process of teaching and learning using autobiography and repertory grid. *Review of Educational Research, 62* (2), 205-225.

Stone, D. N. (1990). Assumptions and values in the practice of information systems evaluation. *Journal of Information Systems, 4* (2), 1-17.

Taşdemir, A. ve Tay, B. (2007). Fen bilgisi öğretiminde öğrencilerin öğrenme stratejilerini kullanmalarının akademik başarıya etkileri. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi, 20* (1), 173-187.

Ünal, G. ve Ergin, Ö. (2006). Buluş yoluyla fen öğretiminin öğrencilerin akademik başarılarına, öğrenme yaklaşımlarına ve tutumlarına etkisi. *Türk Fen Eğitimi Dergisi, 3* (1), 36-52.

White, R. T., & Gunstone, R. F. (1992). *Probing understanding.* Great Britain: Falmer Press.

Yener, N. (2006). *İlköğretim 7.sınıf fen bilgisi dersinde çevremizde hangi ekosistemler var ve buralarda neler oluyor? Konusunun kavram haritalarıyla işlenmesinin öğrenci başarı ve tutumu üzerindeki etkisi.* Yayınlanmamış yüksek lisans tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Yıldız-Duban, N. (2006). The effect of teaching learning strategies on the academic success of the 5th grade primary school students in the science course. *Eurasian Journal of Educational Research, 22,* 111-120.

Yılmaz, H. (1997). *Eğitimde ölçme ve değerlendirme* (2. bs.). Konya: Öz Eğitim Basım Yayın.

Yılmaz, H., & Huyugüzel-Çavaş, P. (2006). The effect of the 4-E learning cycle method on students'understanding of electricity. *Journal of Turkish Science Education, 3* (1), 2-18.