The Measurement of Students' Achievement in Teaching Primary School Fifth Year Mathematics Classes*

Ahmet DOĞANAY**, Ayten Pınar BAL***

Abstract

The aim of this study was to investigate students' and teachers' point of views about preparing measurement tools used in mathematics classes, the level of learning that these tools are intended to measure, how often they are used and how they are scored in terms of assessing 5th grade primary school mathematic courses. The population of the study for the quantitative data was 226 primary school fifth-year classroom teachers and 881 students in the central school districts of Adana. Furthermore, in order to obtain qualitative data 25 teachers and 45 students were selected by using criterion sampling. As data collection tools, the Measurement and Evaluation Questionnaire (MEQ) developed by the researchers, semi-structured interview forms, and examination materials were used. Descriptive statistics and content analysis were performed on the data. Also, document analysis was done. As for the results, it was seen that teachers very frequently considered their students' levels and their acquisition when preparing traditional and alternative measurement tools. However, they ignored the analysis-synthesis level. In this context, it can be proposed that while preparing traditional and alternative assessment tools it should be paid attention to measure higher order thinking skills of students and to use more alternative assessment tools.

In recent years, measurement and evaluation have gained importance as a significant factor in effective learning and teaching. This has been reinforced by related studies in the field as well (Black and William, 1998). Regarding mathematics courses, according to the National Council of Teachers of Mathematics (NCTM), measurement and evaluation have been indispensable parts of mathematics teaching as they increase students' interests towards mathematics (NCTM, 2000). Therefore, measurement and evaluation can be regarded as a process of collecting evidence about students' mathematical skills and knowledge in order to find out students' tendency towards mathematics.

Nowadays, as a result of developments in cognitive psychology, traditional approaches based on behaviorism have been replaced by alternative approaches such as structuralism (detailed and multiple), investigating students' individual abilities, high-level thinking skills, revealing their manual skills, integrating new knowledge into already established one in solving complex problems, reasoning in relation to the real life and using different measurement tools in that process (Stiggins, 1999; Sheffield & Cruikshank, 2000; Krulick et al., 2003; Dominguez Carmino, 2004). Alternative measurement and evaluation include all evaluations, excluding traditional measurements (Atkin, Black & Coffey, 2001; Bryant, 2001; Atılgan, 2006; Bahar, Nartgün, Durmuş, & Bıçak, 2006). In addition, Palm (2008) said that alternative measurement and evaluation has begun after 1990, following the criticisms about multiple choice exams in the USA. Also, alternative measurement and evaluation provide students some skills which may help them solve daily and business life problems (Green & Emerson, 2008; Weigold, 1999). According to Wiggins (1989a), the main characteristics of alternative measurement and evaluation are that they are realistic, judicial, and innovative. Moreover, they guide students about what to do and provide opportunities in which students can use their prior knowledge and skills comfortably and evaluate their skills efficiently and productively.

It has been thought that investigating teachers' and students' perspectives about measurement tools in the mathematics teaching program followed since 2005 in Turkey is important as this increases the quality of education. In line with this, the study intends to find out teachers' and students' point of views about the process of preparation of measurement tools in the mathematics course, the level that the measurement tools intend to investigate, the frequency of administration of these

tools, and finally scoring system of these tools. Therefore, this research aims to answer the following questions:

1) How are the measurement and evaluation tools included in the primary school fifth-year mathematics course prepared?

2) Which levels of learning do measurement and evaluation tools prepared in the primary school fifth-year mathematics course aim to measure?

3) How often are measurement and evaluation tools prepared in the primary school fifth year mathematics course used?

4) How are measurement and evaluation tools prepared in the primary school fifth year mathematics course scored?

Method
Research Design and Sample

This research is based on both qualitative and quantitative methods. The population of the study is primary school teachers teaching fifth year students and their students in state schools directed by the Ministry of National Education in the central school districts of Adana. 46.8 % of the participant students are girls and 5.2 % are boys. As for the participating teachers, 59.8 % are females and 40.2 % are males.

Measurement Instruments

In this study, the Measurement and Evaluation Questionnaire (MEQ) developed by the researchers, semi-structured interview forms, and examination materials were used. The quantitative data of the study were collected through the MEQ. For the scope validity of the MEQ prepared for the teachers and students, 10 instructors specialized in mathematics teaching in different universities in Turkey were asked for advice. Moreover, 10 instructors working at Çukurova University, the Education Faculty, the Educational Sciences and the Primary School Teaching Department, 2 primary school teaching inspectors and 10 primary school teachers teaching primary school fifth-year students were asked for advice. Then, following the suggestions, the MEQ teacher and student forms were revised and finalized. Next, Cronbach Alpha as a reliability coefficient of the MEQ teacher form was calculated and found as .78, .79 for the preparation of the measurement tools, .74, .79

for the learning level that the tools aimed to measure, .88 for the use of frequency of the measurement tools, .54, .80 for the scoring of the measurement tools respectively.

As all these scores are higher than .70, it can be concluded that the MEQ is reliable. As for the MEQ student form, the Cronbach Alpha reliability was found .81. Furthermore, the students and teachers were interviewed about the measurement tools. The measurement tools were analyzed in line with these interview results. When preparing the interview forms, the theoretical knowledge in the related literature, data collection tools in the parallel studies, the parts in the questionnaire and the results of the pre-interviews and the experts' point of views were considered as well. The semi-structured interview forms were reviewed in line with the recommendations by ten instructors at Çukurova University, the Faculty of Education, the Educational Sciences and Primary School Teaching Department.

Then, the questions were administered to two volunteer teachers and three students as a pilot study in order to test whether the questions were comprehensible and applicable. No problems were encountered in this process.

Regarding the document analysis, the copies of exam papers that the teachers gave in the 2006-2007 academic year were also collected. In addition, the students' works and some photocopies of the students' class portfolios were taken as examples or their photographs were taken within the scope of alternative measurement and evaluation.

Data Analysis

For descriptive statistics, SPSS-Windows 13.0 was used to analyze the quantitative data of the study. Content analysis was performed for the qualitative data. In content analysis, first, the main concepts are identified in the collected data. Then, these concepts are organized in a logical order and appropriate themes are investigated (Yıldırım & Şimşek, 1999). While deciding on the codes, the data based on the teachers' and the students' interview forms were read line by line. Next, both already established codes and newly emerged ones were identified and highlighted on the data. Following, similarities and differences among the codes were taken into account and the data were re-grouped accordingly. Lastly, thematic coding was performed on the related codes

(Yıldırım & Şimşek, 1999). While performing the thematic coding, meaningful groups were formed in order not to lead to mismatch or confusion in other codes. These codes were presented to the specialists who worked about program development and mathematics teaching. They were revised in line with the specialists' recommendations.

Different labels were given to the first interviewed teacher as T1, the second teacher as T2, and similarly to the first interviewed student as S1. In the study, the document analysis was used to reinforce the quantitative data and to provide alternative explanations to the results of the study. Firstly, the exam papers given in the 2006-2007 academic year were photocopied and copies of assignments were taken. Based on this document analysis, supportive and alternative explanations were provided in line with the aims of the study.

Results

The findings based on the teacher input are given below in relation to sub-aims of the study. As a first sub-aim of the research, teachers mentioned that they frequently considered their students' levels ($\overline{\text{x}}$=4.61), their students' acquisition ($\overline{\text{x}}$= 4.50) and the properties of the quality to be measured ($\overline{\text{x}}$= 4.40) in preparing traditional measurement tools. Then, the items that the teachers took into account at a very low level while preparing traditional measurement tools were "*I prepare the tools in collaboration with my students*" ($\overline{\text{x}}$=3.39) and "*I prepare the tools with the fifth-year class teachers*" ($\overline{\text{x}}$=3.56). Also, in the interviews, the teachers said that they gave importance to students' qualifications, question characteristics and acquisition.

As for the preparation of the alternative measurement tools, the items that the teachers most frequently followed were; according to their mean; "*I consider the students' level while preparing*" ($\overline{\text{x}}$= 4.53), "*I consider the acquisitions while preparing*" ($\overline{\text{x}}$= 4.42) and "*I consider the properties of the quality to be measured*" ($\overline{\text{x}}$=4.36). As for the items that were considered at the lowest level, they were "*I prepare the tools individually*" ($\overline{\text{x}}$=3.65), "*I prepare the tools in collaboration with the students*" ($\overline{\text{x}}$=3.75) and "*I prepare the tools with the fifth-year class teachers*" ($\overline{\text{x}}$=3.75). Also, in the interviews, the teachers mentioned that they took into account students' qualifications, acquisition and the material while preparing the alternative measurement tools.

The second sub-aim of the study is about the learning level that traditional and alternative measurement tools prepared in the mathematics classes aimed to measure. Regarding this, the teachers said that they frequently paid attention to *remembering* ($\overline{X}$=4.20), *problem solving* ($\overline{X}$=4.20), *comprehension* ($\overline{X}$=4.17). The items that were rarely considered by the teachers were *analysis-synthesis* ($\overline{X}$=3.78), and "*it only requires processing*" *(*$\overline{X}$=3.89). According to the results about the learning level that alternative measurement tools aimed to measure, the most frequently considered choices were "*application*" ($\overline{X}$=4.14), "*comprehension*" ($\overline{X}$=4.07) and "*requirement of problem solving*" ($\overline{X}$=3.99). The least frequently considered choices about alternative measurement tools were "*it only requires processing*" ($\overline{X}$=3.86), "*analysis-synthesis*" ($\overline{X}$=3.88).

In addition to the findings given above, all exam papers, performance assignments, student portfolios and projects tasks in the 2006-2007 academic year were taken from the schools. 73 exam papers were analyzed within the scope of traditional measurement and evaluation. 40.91 % of the multiple choice questions in the exam papers required problem solving and 24.24 % of them only required operation. Besides this, 13.64 % of the multiple choice questions were at remembering level and 21.21 % of them were at comprehension level. As for the written exam questions, 96.85 % of them were problem solving questions and 3.15 % of them only required operation.

In addition to these findings, rubric analysis was conducted on the students' performance assignments, project tasks and portfolios within the scope of alternative measurement and evaluation tools. Five rubrics were identified in collaboration with 4 specialists and through the related research in the field. The rubrics were "Data Collection", "Data Arrangement", "Data Application", "Data Interpretation" and "Originality". 53 performance assignments were evaluated in line with these five rubrics. According to this, it can be said that the criteria on "Data Application" ($\overline{X}$=2.09), and "Data Collection" ($\overline{X}$=1.94) and "Data Arrangement" ($\overline{X}$=1.94) were achieved. However, the criterion "Originality" ($\overline{X}$=1.60) was not fully achieved.

The most frequently used measurement tools by the teachers "twice in two week-period" were observation (38.0 %) and short-answer questions (30.8 %). Multiple choice exams were given at the end of each unit (35.4 %). The least frequently used measurement tools "once in a term" were project assignments (78.5 %) and performance tasks (50.9

%). As in the teachers' questionnaires, "*observation*" was said to be the most frequently used measurement tool in the students' questionnaire results. The students mentioned that observation technique was used "once or twice in a week" (64.8 %). Also, according to the students' point of views, the most frequently used measurement tools -once or twice in a week- were "*short-answer questions*" (61.6 %), "*interview*" (66 %), "*multiple choice questions*" (53.1 %). In addition to these findings, according to the students' point of views, the least frequently used measurement tool "never" was "the Attitude Inventory" (57.5 %). Other least frequently used measurement tools -once in a term" were "project assignments" (46.2 %) and "*performance tasks*" (36.7 %).

As for the scoring system in alternative and traditional measurement tools, the teachers frequently preferred the choices as in the following: "*I consider the difficulty level of questions*" ($\overline{\text{X}}$=4.39), "*I consider the students' knowledge level*" ($\overline{\text{X}}$=4.20) and "*I use a detailed answer key*" ($\overline{\text{X}}$=4.11). Also, when scoring the traditional measurement tools, the teachers preferred the following choices least frequently: "*I only give a mark to the result in written exams*" ($\overline{\text{X}}$=2.58) and "*I use a shared answer key with the fifth-year class teachers*".

When the teachers scored the alternative measurement tools in the mathematics classes, they said that the most frequently used methods were according to the mean: "*I consider that my students reveal their individual skills*" ($\overline{\text{X}}$=4.46), "*I consider my students' daily performances*" ($\overline{\text{X}}$=4.30) and "*I consider my students' knowledge level*" ($\overline{\text{X}}$=4.29). On the other hand, the least frequently preferred choices by the teachers were "*I use rubrics*" ($\overline{\text{X}}$=3.70),

"*I use a detailed answer key in scoring performance tasks*" ($\overline{\text{X}}$=3.70) and "*I use a detailed answer key in scoring project assignments*" ($\overline{\text{X}}$=3.77).

According to the interviews with the teachers, the first theme that the teachers gave importance in scoring traditional measurement tools is questions. 19 of the teachers said that they gave a mark to the way that the student followed in solving the problem; in other words, they said that they assigned a mark to each correct phase in the answers. Also, the teachers were asked about which points they took into consideration while scoring alternative measurement tools in the interviews. They were first asked whether they used rubrics or not in evaluating alternative measurement tools. 19 of the teachers said that they used rubrics but 6 of them said that they did not use rubrics. 10 of the teachers using

rubrics explained that they gave attention to class level when preparing rubrics and 9 of them said that they followed the rubrics in the guide books. Lastly, as for the students' interview results, four themes emerged from the scoring system of performance tasks and project assignments: content, report writing, visuality and presentation.

Discussion

In this study, it was aimed to investigate teachers' and students' point of views about how the measurement tools were prepared in the mathematics classes, which learning level the measurement tools measured, how often they were used and how they were scored through both qualitative and quantitative method in our country.

When questionnaire and interview results were integrated, it was clearly seen that teachers considered their students' level most frequently in preparing traditional and alternative measurement tools. This finding is in line with the results by Senk et al. (1997a). Senk et al. (1997) highlighted that the class tests should be prepared in a way that they evaluate students' skills in their study. Similar to this, Hopkins (1999) mentioned that alternative measurement and evaluation tools should be used in a way that they emphasize students' characteristics. According to NCTM (2000), students' ages, experiences, and interests should be given importance in choosing measurement tools. In addition to this, the teachers secondarily considered the acquisition in preparing traditional and alternative measurement tools. The Ministry of National Education (2005) also mentioned the importance of acquisition when talking about measurement and evaluation tools.

In this research, regarding the learning level that traditional measurement tools intended to evaluate, the MEQ teachers' form and document analysis were used. According to the questionnaire results, the teachers mostly cared for problem solving, remembering and comprehension in preparing their exams. Similarly, the document analysis indicated that the teachers gave importance to problem solving, comprehension and operating during exam preparation phase. Therefore, it can be said that the questionnaire and document analysis findings support each other.

Most of the teachers told that they prepared the exam questions which highlighted problem solving and operating. These findings are partially parallel to research results by Saxe et al. (1997), Miller (2004), Archbald and Grant (2000) who conducted studies with the mathematics teachers as well.

Then, the learning level that alternative measurement tools intended to measure was focused through the questionnaire results. According to this, it was seen that the teachers firstly preferred application level and then comprehension level. For example, the teacher (T5) explained that "*we assigned the students project and performance assignments so that they could develop their manual skills. We did not give written documents. While preparing these assignments, we wanted that the students' daily activities were revealed*".

As for the frequency of use of measurement tools, it was found that the questionnaire results by the students and the teachers supported each other. In both questionnaires, the most frequent measurement tools were "observation", and "short answer questions". Then, it was said that "observation" was used once or twice in two-week period by the teachers and students. However, according to the interview findings, although most of the teachers said that they observed, they did not use the observation as a measurement tool. Therefore, it can be said that the teachers regarded their class observation as a measurement tool. In line with this, a teacher said like this:

> "… I make mini revisions after I complete the topic… For example, when there are five minutes left after I finish the topic. I ask two translation questions the most. "I say that you should do it on your own, do not cheat" Then, I walk around the class. I observe the students. I do not count. I keep them in my mind. We evaluate the students based on this. I know the student since he was first-year student …" (T14).

According to the teachers' questionnaire, "short-answer questions" and "multiple choice questions" were asked once a month, whereas, according to the students' questionnaire, these measurement tools were used once or twice a week. There is a considerable inconsistency at that point between the teachers' and the students' answers. This may have derived from the fact that the students considered homework assignments as measurement tools.

On other hand, the interview findings showed that among the traditional measurement tools, the most frequently administrated measurement tools were "mixed exams" (written exams, multiple choice exams, true-false questions, and fill in the blanks questions) and written exams and these were said to be given at the end of units. These findings support the teachers' and the students' findings.

In line with these results, the qualitative and quantitative findings support each other about the frequency of use of the measurement tools. Mixed exams, short-answer exams and multiple choice exams and written exams were most frequently used measurement tools. This is parallel to studies Saxe et al. (1997), Senk et al. (1997a), Pilten (2001), Miller (2004), Çakan (2004), Watt (2005), Susuwele-Banda (2005), Kalender (2006), Birgin (2007), Erdemir (2007), Sarıer (2007), Erdal (2007), Özdaş et al. (2005) and Güven and Eskitürk (2007).

It was also concluded that project and performance assignments were given once a term or once in a unit according to the questionnaire and interview results. This is in line with the study by Yurday (2006). This stems from the necessity that the Ministry of National Education notice, at least one project in a year and at least one performance task in a term must be assigned.

The results indicated that the teachers scored the questions according to the difficulty level of the questions. This is consistent with the finding by Erdemir (2007). In his study, he investigated how efficient primary school teachers' used measurement and evaluation tools and he came into the similar results. Also, another important finding was that the teachers gave a mark to the way that the students followed in the operation. This is in line with the study conducted by Van Den Heuvel-Panhuizen and Fosnot (1996). Moreover, Senk et al. (1997a) presented the similar findings conducted in secondary school mathematics classes. However, these results contradict with the findings by Kalender (2006). In his study with primary school teachers in İzmir, the participants said that they did not give a score to the way that the students followed. Instead, they scored the correct result. Because the teachers thought that the mathematics is a positive science and there is only one correct solution. Nevertheless, the main philosophy of alternative measurement and evaluation tools are the process, not the product.

In scoring the alternative measurement and evaluation tools, the questionnaire findings by the teachers and interview results overlapped.

The questionnaire results showed that the students' individual skills, daily performances and their knowledge level were considered in scoring alternative measurement and evaluation tools, whereas, the interview results indicated that both the task, itself, and the students' level were taken into account. It can be inferred that teachers are aware of the

objectives of the alternative measurement and evaluation tools and they reflect this on their evaluation system. This inference is parallel to the finding by Smith (2003).

Similarly, it was found that rubric was used by most of the teachers (19/25) according to the interview results with the teachers in alternative measurement tools. The questionnaire results also supported this finding. However, this contradicts with the results by Saxe et al. (1997). They analyzed primary and secondary school teachers' perceptions about the revised mathematics program. According to the results of this study, mathematics teachers used rubric as a new measurement method unconsciously and at a very low level. Similarly, Lim and Colgan (2005) conducted a study in ninth-class mathematics classes. They found that rubric was rarely used and some difficulties were encountered in the process.

The teachers mentioned in their interviews that most of the teachers evaluated according to the task, itself. Also, the students' interviews indicated the focus on the task. Parallel to this, it can be said that both the teachers' and the students' interview results overlapped and objectivity principle was considered in scoring alternative measurement and evaluation tools. Similarly, Long (2001) also dealt with objectivity in mathematics classes and said that the evaluation was done in line with this principle.

As a result, it was seen that teachers took into account students' level, acquisition, and the characteristics of the quality to be measured in preparing alternative and traditional measurement and evaluation tools. Problem solving and comprehension were considered about the learning level that the evaluation tools intended to measure. The analysis-synthesis level was the point that was considered at the lowest level in the preparation of these measurement tools. As for the frequency of the measurement tools, the most frequently used ones were "observation", "short-answer questions", and "multiple choice questions". It was also observed that alternative measurement and evaluation tools were used less frequent than traditional measurement tools. In scoring traditional measurement tools, the teachers told that they mostly considered the content of the answers and the students' knowledge level, whereas, in alternative measurement and evaluation tools scoring, class participation and the task assigned were given importance. In short, it can be concluded that it is important to prepare alternative and traditional measurement and evaluation tools which reveal students' high level thinking skills. Also, alternative measurement tools should be used more often.

References/Kaynakça

Archbald, D. A., & Grant, T. J. (2000). What's on the test? An analytical framework and findings from an examination of teachers' math tests. *Educatıonal Assessment*, *6*(4), 221-256.

Atılgan, H. (2006). Değerlendirme ve not verme. H. Atılgan, (Edt.), *Eğitimde ölçme ve değerlendirme* içinde (s. 405-454). Ankara: Anı Yayıncılık.

Atkin, J. M., Black, P., & Coffey, J. (2001). *Classroom assessment and the national science education standards.* Washington, DC: National Academies Press.

Aydın, F. (2005). Öğretmenlerin alternatif ölçme değerlendirme konusundaki düşünceleri ve uyguladıkları. H. Kıran (Edt.), *XIV. ulusal eğitim bilimleri kongresi* kitabı içinde (c. 1, s. 775-779). Ankara: Anı Yayıncılık.

Bachman, L. F. (2002). Alternative interpretations of alternative assessments: some validity issues in educational performance assessments. *Educational Measurement: Issues and Practice*, *21*(3), 5-18.

Bahar, M., Nartgün, Z., Durmuş, S. & Bıçak, B. (2006). *Geleneksel ve alternatif ölçme ve değerlendirme öğretmen el kitabı*. Ankara: PegemA.

Baki, A. & Birgin, O. (2004). Alternatif bir değerlendirme aracı olarak bilgisayar destekli bireysel gelişim dosyası uygulamasından yansımalar. *The Turkish Online Journal Of Educational Technology, 3*(3), 11. www.tojet.net/articles/3311.htm adresinden 8 Mart 2005 tarihinde indirilmiştir.

Birgin O. (2007). Sınıf öğretmeni adaylarının ölçme ve değerlendirme konusundaki okur-yazarlık düzeylerinin incelenmesi. E. Erginer (Edt.), *XVI. ulusal eğitim bilimleri kongresi* kitabı içinde (c. 3, s. 498-503). Ankara: Detay Yayıncılık.

Black, P. & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-148.

Bryant, D. D. (2001). *The perception of secondary mathematics teachers in Christian schools on the effectiveness of alternative assessment on academic achievement.* Unpublished master's thesis, University of Memphis, Memphis.

Burke, K. (1999). *How to assess authentic learning (3rd Ed),* Arlington Heights, Illinois: Skyligth Professional Development.

Burrill, J., Feijs, E. , Meyer, M., Reeuwijk, M. V., Webb, D., & Wijers, M. (2001). *The role of assessment standards based middle school mathematics curriculum materials.* Retrieved December 8, 2008 from www.showmecenter.missouri.edu

Carnevale, J. (2006). *The Impact of self-assessment on mathematics teachers' beliefs and reform practices.* Unpublished master thesis, University of Toronto Ontario, Canada.

Charlesworth, R., & Lind, K. K. (2003). *Math and science for young children* (4th Ed.). Clifton Park, NY: Thomson Delmar Learning.

Cooney, T. J., Sanchez, W. B., & Ice, N. F. (2001). Interpreting teachers' movement toward reform in mathematics. *The Mathematics Educator*, *11*(1), 10-14.

Creswell, J. W. (1998). *Qualitative and inquiry and research design choosing among five tradition*s. Thousand Oaks: Sage Publications.

Creswell, J. W. (2003). *Research design qualitative and quantitative and mixed methods approaches.* Thousand Oaks: Sage Publications.

Çakan, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: İlk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 37*(2), 99-114.

Dominguez Carmino, G. (2004). *Designing an assessment tool to describe students' mathematics knowledge*. Unpublished doctoral dissertation, Purdiee University West Lafayete, Indiana, USA.

Eisner, E. W. (1999). The uses and limits of performance assessment. *Phi Delta Kappan*, *80*(9), 658-660.

Erdal, H. (2007). 2005 *İlköğretim matematik programı ölçme değerlendirme kısmının incelenmesi (Afyonkarahisar ili örneği)*. Yayımlanmamış yüksek lisans tezi, Afyon Kocatepe Üniversitesi Sosyal Bilimler Enstitüsü, Afyonkarahisar.

Erdemir, Z. A. (2007). *İlköğretim ikinci kademe öğretmenlerinin ölçme ve değerlendirme tekniklerini etkin kullanabilme yeterliliklerinin araştırılması (Kahramanmaraş örneği)*. Yayımlanmamış yüksek lisans tezi, Kahramanmaraş Sütçü İmam Üniversitesi Sosyal Bilimler Enstitüsü, Kahramanmaraş.

Gelbal, S. & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *33*, 135-145.

Gözütok, D. Akgün, Ö. E. & Karacaoğlu C. (2005). İlköğretim programlarının öğretmen yeterlikleri açısından değerlendirilmesi. M.Işık Tekışık Tanış (Edit.) *Eğitimde yansımalar: VIII Yeni ilköğretim programlarını değerlendirme sempozyumu bildiriler kitabı*, 14-16 Kasım 2005. (s. 17-40). Ankara: Sim Matbaası.

Green, K., & Emerson, A. (2008). Reorganizing freshman business mathematics II: Authentic assessment in mathematics through professional memos. *Teaching Mathematics and ıts Applications, 27*(2), 66-80.

Güven, B. & Eskitürk, M. (2007). Sınıf öğretmenlerinin ölçme ve değerlendirmede kullandıkları yöntem ve teknikler. E. Erginer (Edt.), *XVI. ulusal eğitim bilimleri kongresi* kitabı içinde (c. 3, s. 504-511) Ankara: Detay Yayıncılık.

Haertel E. H. (1999). Performance assessment and education reform. *Phi Kappan Delta, 80*(9), 662-666.

Heddens, J. W., & Speer, W. R. (2006). *Today's MATHEMATICS: Concepts, methods and instructional activities* (11th Ed.). Hoboken NJ: John Wiley and Sons, Inc.

Holaway-Johnson, C. A. (2005). *Best practices in middle school mathematics*. Unpublished doctoral dissertation, University of Arkansas, Arkansas.

Hopkins, M. H. (1999). Practicing what we preach: Authentic assesment in mathematics. *Assessment of Effective Intervention*, 25(1), 15-30.

Jimarez, T. (2005). *Does alignment of constructivist teaching, curriculum, and assessment strategies promote meaningful learning?* Unpublished doctoral dissertation, New Mexico State University Las Crues, New Mexico.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26.

Kalender, A. (2006). *Sınıf öğretmenlerinin yapılandırmacı yaklaşım temelli "yeni matematik programı"nın uygulanması sürecinde karşılaştığı sorunlar ve bu sorunların çözümüne yönelik önerileri*. Yayımlanmamış yüksek lisans tezi, Dokuz Eylül Üniversitesi Eğitim Bilimleri Enstitüsü, İzmir.

Karaca, E. (2003). *Öğretmen adaylarının ölçme ve değerlendirme yeterliliklerine ilişkin algıları*. Yayımlanmamış doktora tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Karasar, N. (1999). *Bilimsel araştırma yöntemleri* (9. baskı). Ankara: Nobel Dağıtım.

Kartallıoğlu, F. (2005). *Yeni ilköğretim programlarının uygulandığı pilot okullardaki öğretmenlerin yeni program ve pilot çalışmalar hakkındaki görüşleri.* Yayımlanmamış yüksek lisans tezi, Abant İzzet baysal Üniversitesi Sosyal Bilimler Enstitüsü, Bolu.

Keller-Cogan, M. M. (1996). *Student voices: High school students perceptions of instructional and assessment strategies in traditional and alternative settings.* Unpublished master's thesis, University of Rochester, Rochester, New York.

Krulick, S., Rudnick, J., & Milou, E. (2003). *Teaching mathematics in the middle school.* New York: Pearson Education.

Kulm, G. (1993). *A theory of classroom assessment and teacher practice in mathematics.* Retrieved March 2, 2006 from http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/29/a6/c6.pdf

Kulm, G. (1994). *Mathematics assessment: What works in the classroom.* San Francisco: Jossey-Bass Publishers.

Kutlu, Ö. (2005). Yeni ilköğretim programlarının 'öğrenci başarısındaki gelişimi değerlendirme' boyutu açısından incelenmesi. M.Işık Tekışık Tanış (Edit.) *Eğitimde yansımalar: VIII Yeni ilköğretim programlarını değerlendirme sempozyumu bildiriler kitabı*, 14-16 Kasım 2005. (s. 64-71). Ankara: Sim Matbaası.

Liebers, C. (1999). Journals and portfolios: Alternative assessment for pre service teachers. *Teaching Children Mathematics*, *6*(3), 164-169.

Lim, L. (2002). Implementing multiple assessments in a grade 9 applied mathematics class: A case study of one teacher and his students, *Master Thesis*, Queen's University Kingston Ontario, Canada.

Lim, L., & Colgan, L. (2005). Implementing multiple assessments in mathematics: An Action research study of one teacher and his students. *The Ontario Action Researcher*, 7(1), 1-6. Retrieved January 8, 2006 from http://www.nipissingu.ca/oar/archive-Vol7No1-V713E.htm

Long, V. (2001). The myth of objectivity in mathematics assessment. *Mathematics Teacher, 94*(1), 31-37.

Mabry, L. (1999). *Writing to the rubric: Lingering effect of traditional standardized testing ob direct writing assessment.* Retrieved May 28, 2007 from http://www.pdkintl.org/kappan/kmab9905.htm

Maxwell V. L., & Lassak, M. B. (2008). An experiment in using portfolio in the middle school. *Mathematics Teaching in the Middle School*, *13*(7), 404-409.

MEB, (2005). *İlköğretim matematik dersi (1-5. sınıflar) öğretimi programı*. Ankara: Devlet Kitapları Müdürlüğü Basımevi.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd Ed.). Thousand Oaks: Sage Publications.

Miller, T. (2004). *Assessment in practice grade 9 academic and applied mathematics*. Unpublished master's thesis, Queen's University, Kingston, Ontario, Canada.

Moskal, B. M. (2000). An assessment model for the mathematics classroom. *Mathematics Teaching in the Middle School*, *6*(3), 192-194.

Motsoeneng, K. G. (2005). *The attitude of teacher and parents and learners involved in primary and intermediate schools in the Thabı mofutsanyana district regarding assesment reform in education.* Unpublished master's thesis, Bloemfontein University, Mofutsanya Thabo.

Mulvenon, S. W., McKenzie, S. C., Connors, J. V., & Williams, T. L. (2003). Teachers' attitudes toward the use of standardized testing: Implications for practice. *Arkansas Educational Research and Policy Studies Journal*, *3*(1), 61-80.

Myers, S. (2008). Authentic assessment. *Research Starters Education: Authentic Assessment, 1* (1), 1-12. Retrieved January 12, 2006 from http://search.ebscohost.com/login.aspx?direct=true&db=e0h&AN=27577900&site=ehost-live

Nash, L. E. (1993). *What they know vs. what they show: An investigation of teachers' practices and perceptions regarding student assessments.* Unpublished doctoral dissertation, Georgia State University Georgia, United States.

NCTM., (1995). *Assessments standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

NCTM., (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Olkun, S. (2006). *Yeni öğretim programlarını inceleme ve değerlendirme raporu*. http://ilkogretim-online.org.tr/vol5say1/yenimufredat_raporu[1].pdf adresinden 11 Aralık 2007 tarihinde indirilmiştir.

Ostrow J. (1999). *Making problems creating solution challenging young mathematicians,* Portland, Maine: Stenhouse Publishers.

Özdaş, A. , Tanışlı, D. , Köse, N. Y. & Kılıç, Ç. (2005). Yeni ilköğretim matematik dersi (1.-5. sınıflar) öğretim programının öğretmen görüşlerine dayalı olarak değerlendirilmesi. M.Işık Tekışık Tanış (Edit.) *Eğitimde yansımalar: VIII Yeni ilköğretim programlarını değerlendirme sempozyumu bildiriler kitabı*, 14-16 Kasım 2005. (s. 239-255). Ankara: Sim Matbaası.

Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Proctical Assessment Research & Evaluation, 13*(4), 1-11.

Pandey, T., & Smith, T. R. (1991). *A sampler of mathematics assessment* (Report No. ISBN-0-8011-0972-8). (ERIC Document Reproduction Service No. ED341553). Retrieved March 30, 2006 from www.eric.com.

Payne, J. N. (1993). *Mathematics for the young child* (2nd Ed.). Reston, Virginia: National Council of Teachers of Mathematics.

Pilten, P. (2001). *İlköğretim sınıf öğretmenlerinin ölçme ve değerlendirme alanındaki anlayış ve uygulamalarının değerlendirilmesi.* Yayımlanmamış yüksek lisans tezi, Selçuk Üniversitesi Sosyal Bilimler Enstitüsü, Konya.

Romberg, T. A. & Lange, J. (2005). Research in assessment practices. In T. A. Romberg, T. P. Carpenter ve F. Dremock (Ed.), *Understanding mathematics and science matters* (pp.279-307). Mahway, New Jersey: Lawrence Erlbaum Assosiates Publishers.

Romberg, T. A., Carpenter, T. P., & Kwako, J. (2005). Standart based reform and teaching for understanding. In T. A. Romberg, T. P. Carpenter & F. Dremock (Ed.), *Understanding mathematics and science matters* (pp. 3-26). Mahwah, New Jersey: Lawrence Erlbaum Assosiates Publishers.

Sarıer, Y. (2007). *Altıncı sınıf matematik öğretmenlerinin matematik dersi öğretim programına ilişkin görüşleri.* Yayımlanmamış yüksek lisans tezi, Eskişehir Osmangazi Üniveristesi Fen Bilimleri Enstitüsü, Eskişehir.

Saxe, G. B. , Franke, M. L., Gearhart, M., Howard, S. & Crockett, M. (1997). *'Teachers' shifting assessment practices in the context of educational reform in mathematics.* (Tech. Rep. No. 471) CRESST University of California, Los Angeles. Retrieved November 12, 2006 from www.cresst.org.

Senk, S. L., Beckmann C. E., & Thompson, D. R. (1997a). Assessment and grading in high school mathematics classroom. *Journal of Research in Mathematics Education*, *28*(2), 187-215.

Senk, S. L., Beckmann C. E. & Thompson, D. R. (1997b). Improving classroom tests as a means of improving assessment. *The Mathematics Teacher, 90*(1), 58-64.

Sheffield, L. J. & Cruikshank, D. E. (2000). *Teaching and learning elementary and middle school mathematics.* New York: John Wiley and Sons.

Shepard, L. A. (2001). The role of classroom assessment in teaching and learning, CSE Technical Report, U.S.A: California; 2000-02-00 http://www.cse.ucla.edu/products/Reports/TECH517.pdf adresinden 12 Ocak 2006 tarihinde alınmıştır

Smith S. S. (2003). *Early chilhood mathematics (3rd Ed.*), Newyork: Pearson Education.

Solomon, P. G. (2003). *The curriculum bridge: From standards to actual classroom practice* (2nd Ed.). California: Corwin Press.

Stiggins, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan, 83*(3), 191-198.

Stiggins, R. J. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan, 87*(4), 324-328.

Susuwele-Banda, W. J. (2005). *Classroom assessment in Malawi: Teachers' perceptions and practices in mathematics.* Unpublished doctoral dissertation, Blackburg Virginia Polytechnic Institute, Virginia.

Van De Walle, J. A. (2001). *Elementary and middle school mathematics: Teaching developmentally.* Boston: Allyn and Bacon.

Van Den Heuvel-Panhuizen, M. & Fosnot, C. T. (1996). Assessment of mathematics achievements: Not only the answers count. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th conference of the international group for the psychology of mathematics education* (Vol. 4, pp. 335-342). Utrecht, The Netherlands: Freudenthal Institute, Utrecht University.

Watt, H. M. G. (2005). Attitudes to the use of alternative assessment methods in mathematics: A study with secondary mathematics teacher in Sydney, Australia. *Educational Studies in Mathematics*, *58*, 21-44.

Webb, D. C. (1992). Assessment of student' knowledge of mathematics: Steps toward a theory. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 661-683). New York: Macmillan Library.

Webb, D. C. (2001). *Instructionally embedded assessment practices of two middle grades mathematics teachers.* Unpublished doctoral dissertation, University of Wisconsin, Madison.

Weigold, J. K. (1999). *Self concept and attitude towards tradional or alternative assessments: An exploration of gender differences in mathematics and science.* Unpublished master's thesis, Eastern Michigan University Ypsilanti, Michigan.

Wiggins, G. P. (1989a). Teaching to the (authentic) test. *Educational Leadership*, *46*(7), 141-147.

Wiggins, G. P. (1989b). A true test. Toward more authentic and equitable assessment. *Phi Delta Kappan*, *70*, 703-713.

Yaşar, Ş., Gültekin, M., Türkkan, B., Yıldız, N. & Girmen, P. (2005). Yeni ilk programlarının uygulanmasına ilişkin sınıf öğretmenlerinin hazır bulunuşluk düzeylerinin ve eğitim gereksinimlerinin belirlenmesi (Eskişehir ili örneği). *Eğitimde yansımalar: VIII. yeni ilköğretim programini değerlendirme sempozyumu* kitabı içinde (s. 51-63). Ankara: Sim Matbaası.

Yıldırım A. & Şimşek H. (1999). *Sosyal bilimlerde nitel araştirma yöntemleri*. Ankara: Seçkin Yayınevi.

Yılmaz, T. (2006). Yenilenen 5. sınıf matematik programı hakkında öğretmen görüşleri (Sakarya ili örneği), *Yüksek Lisans Tezi,* Sakarya Üniversitesi Sosyal Bilimler Enstitüsü, Sakarya.

Yurday, H. (2006). *Lise matematik öğretmenlerinin yeni öğretim programına yaklaşımları*. Yayımlanmamış yüksek lisans tezi, Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü, Trabzon.