# The Effects of Activities and Approachments Intended Performance Improvement on the Students' Performances in Mathematics 

Seval Deniz KILIÇ ${ }^{1}$ \& Hüseyin ALKAN $^{2}$<br>${ }^{1}$ Faculty of Education, Mugla Sttk1 Kocman University, Mugla, Turkey<br>${ }^{2}$ Faculty of Education, Dokuz Eylül University, İzmir, Turkey<br>Correspondence: Faculty of Education, Mugla Sitkı Kocman University, Mugla, Turkey. E-mail: sdenizkilic@mu.edu.tr

Received: June 1, $2018 \quad$ Accepted: July 7, $2018 \quad$ Online Published: July 20, 2018
doi:10.5539/jel.v7n5p146
URL: https://doi.org/10.5539/jel.v7n5p146


#### Abstract

The study presented aims at searching the relation between the learning environment and students' performances in Mathematics. Data of the 27 week-long research were compiled from 9th grade students at an Anatolian High School. Semi-experimental research modelled pre-test post-test control grouped experimental model was used in the study. Before the application, students' gaps regarding the subjects were found out and students were supported by it and their initial performances were evaluated by using the developed non-routine problem. During the application, the learning process continued taking the performance variables into consideration in the experiment group whereas the constructivist learning approach was used in the control group in accordance with the mathematics program. Data gained from students' performance evaluations at the beginning and end of the term, were analysed in order to find out the similarities and differences between the two groups. The findings reveal that there is no meaningful difference in the performance grades of students at the beginning of the term. On the other hand, the average performance grades of the students after the application process show a statistically meaningful difference in favour of the experiment group. Comparisons were made in the experiment group before and after the application.


Keywords: performance, performance assessment, performance in math

## 1. Introduction

One of the main aims of education is to make people grow up as successful individuals in their lives. In the education world, the performance of students can predict their academic achievement because performance is an indicator of students' theoretical knowledge, skills necessary to make use of this knowledge and individual differences. The level of performance acquired and developed throughout the education life can have important repercussions on the business life. Thus, the main indicator of the inseparability of the education process and insert the process is the performance displayed. In short, the level of performance in professional life relies on continuous improvement of the behaviors acquired and developed during education life.
Seen from this perspective, the performance in the professional life is not much different from the performance in the education life. On the other hand, some terms are replaced with new concepts related to the field of profession. For example, in the field of education, the most important aspect of performance is students' knowing some certain knowledge and theories and demonstrating individual skills that need to be exhibited while this knowledge is used. In the process of demonstrating the acquired skills and knowledge, students are expected to utilize their unique thoughts and creativity. That is, while students are expected to have acquired standard knowledge and behaviors, they are also expected to have complete knowledge of the application of methods and techniques. More importantly, the student's exhibiting scientific knowledge by using his/her self-confidence in the application is one of the main indicators of the performance in education.
In a study, it was stated that performance in education is "all the activities exhibited by students in the process of learning and thinking" (Elliott \& Fuchs, 1997). In a similar manner, Kitchen and colleagues define performance as "the approach adopted by the student in the process of thinking" (Kitchen, Cherrington, Gates, Hitchings, Majka, Merk\& Trubow, 2002). There are some other people defining performance as "depth in understanding and high level of competence" (Niemi, 1997). On the other hand, Baker (1997) defines it as "the activities
performed by the student to develop skills and understanding" In a different approach, Wilhite argues that "the student's internalization of the concept is associated with academic performance" (cited in House, Keely, \& Hurst, 1996). In all these definitions, the relationship of performance with understanding is emphasized. In other words, in order to be able to achieve a high level of performance in education, first, it is necessary to understand "what is said", "what is expected" and "what is referred to". That is, one of the indicators of performance is the precise and complete understanding of the event, phenomenon or problem that needs to be addressed.
There are some people defining performance as "the skill of the student's being able to use his/her knowledge related to a subject and to conduct research on this subject" (Slater, 2007). A similar definition is proposed by Sweet. According to him, performance is "the student's skill of using the knowledge he/she has" (Sweet, 1993). As can be seen, in both of the definitions, the application skills of knowledge are emphasized. Thus, another indicator of performance is the application skill.
Performance is also defined as "the student's demonstrating his/her mathematical knowledge, strategies and communication skills" (Spring Branch Independent School District, 2010). Another similar definition is that "performance is the student's level of conceptual learning and methodological application" (Slater, 2007). In these definitions, determination of methods-techniques and strategies comes to the fore.
In the literature, it is suggested that in order to increase performance, individual aptitude, motivation and the effect of environment should be taken into consideration as the predictors of performance (Brecher, 2007). When this approach is adapted to education, it becomes apparent that the learning environment can be a predictor.
Given the delineations above, the safe conclusion is that performance is closely associated with the possession of some certain abilities and skills and their application by the individual. These abilities and skills can be summarized as follows; "conceptual understanding, existence of prior knowledge, problem solving, the skill of being able to explain what has been done, individual differences arising from the education received and the skill of reflecting the content knowledge into application" (Niemi, 1997). When carefully examined, it is seen that with the definitions given in the literature and the conclusions derived from these definitions, the individual's performance level and actual achievement can be defined because "a person's knowing the conceptual field requires the possession of a series of skills to understand, reason and share. Such a series of skills also make the person valued and his/her opinions to be shared and discussed by the community" (Greeno, 1991 cited in Niemi 1997). This is related to the ability to explain what has been done.

When considered within the field of math education, performance "refers to the student's ability of solving problems by using various approaches" (Slater 2007). In fact, "the student's utilizing, synthesizing and evaluating various problem solving strategies and approaches when he/she encounters a problem" is directly related to his/her performance (Shavelson, Baxter \& Pine, 1991; Wiggins, 1989 cited in Slater 2007). In the problem solving process, the student is expected to conclude the solution, explain the stages of solution and prove the accuracy of solution (Baxter, 2001). Moreover, academic performance also involves the explicit demonstration of knowledge, the use of conceptual knowledge for comment and the skill of solving various problems and generalizing knowledge (Niemi, 1997). All of these show that academic performance is related to problem solving. Another argument can be that the performance level of a student best can be seen during a problem solving process because a student who is trying to solve an ambiguous and unusual problem makes use of his/her knowledge most effectively and adds something from himself/herself to the solution with his/her assumptions and interpretations. At that point, the importance of reasoning becomes evident.
In light of the approaches, definitions and explanations made in the literature, the performance indicators are defined as follows in the current study:

- Correctly understanding the event, the phenomenon or the problem.
- Organization of the learning environment
- Possessing a certain level of knowledge about the subject
- Knowing methods and approaches
- Having the application skill
- Having the strategy development skill
- Being able to use individual abilities, and having self-confidence
- Being able to explain what has been done
- Reasoning
- Transferring what has been done to the real life

As a result of this, we created a system including the necessity of compatibility of performance and product or thought with the standards and of their being directed to a certain target (See Figure 1). In the application of the current research, these variables were used as the performance indicators.


Figure 1. Performance system

According to the constructed model, performance in education refers to the student's applying his/her knowledge to different situations, relating it to the daily life, using the time effectively, working cooperatively, coping with the problems he/she encounters, approaching a problem from different perspectives, forming more precise and comprehensive answers and sharing the results obtained effectively (Kılıç, 2011).

## 2. Methodology

### 2.1 Purpose of the Study

The purpose of the current study is to investigate the extent to which students' levels of performance displayed in math classes are affected from activities and approaches targeting performance development. To this end, answers to the following questions were sought:
1- Is there a statistically significant difference between the performance scores of the experimental and control groups before the application?
2- Is there a statistically significant difference between the performance scores of the experimental and control groups after the application?

3- Is there a significant difference between the performance scores of the experimental group taken before and after the application?

### 2.2 Sample and Data Collection

For this experimental study, two classes of $9^{\text {th }}$ grade students from a school located within the city of İzmir in 2010-2011 school year were selected. As the application was conducted in an Anatolian High School, the students were selected according to their SBS exam scores and their scores are close to each other. The families of the students enrolled at the school have similar socio-economic levels. While the classes were formed, a great care was taken to have students from each achievement levels and to keep the number of male and female students equal. Thus, the classes were constructed to be similar to each other. In each class, there were 29 students. In the experimental class, there were 14 female students and 15 male students and in the control class, there were 15 female students and 14 male students.
In the current study, problem solving skills were used to determine the students' performance levels. In the preparation stage of the problems to be used for the measurement of performance, many different problems from
different resources were examined. In this way, many problems were selected and they were solved according to the rubrics developed for each problem and it was discussed whether they comply with the purpose of the study. A group of problems was also created by the authors. To this end, expert opinions were sought and as a result of this expert review, some questions were discarded and some were changed. Initially, it was thought that the tests consisting of $8-10$ questions would be suitable; yet, during the piloting, it was discovered that the students experienced some difficulties in completing the questions on time; thus, the number of the questions in the pretest was reduced to 3 and to 4 in the posttest. Before the actual application, the problems were revised once more and some additions and subtractions were made in line with expert opinions. The 3 questions in the pretest were constructed to probe students' verbal, algebraic and geometric skills. This is because the prior learning of the students in the elementary education can enable them to solve verbal, algebraic and geometric questions. As a warm-up activity, various problems were prepared to comply with the same criteria used in the selection of the questions in the tests and these problems were presented to the students to be solved in the class in line with the stages given in the rubrics; thus, the students were made familiar with what was expected of them. The 4 problems in the posttest were designed to elicit the students' knowledge about clusters, numbers, relations and functions, which are among the subjects studied in the $9^{\text {th }}$ grade curriculum.

### 2.3 Data Analysis

The problems to be solved by the students were evaluated with the aid of a rubric consisting of 4 main stages. While developing this rubric, the samples presented in Marzano, Pickering, \& MCTighe (1993); Woodward, Monroe, \& Baxter (2001) and Lianghuo (2011) were capitalized on. The theoretical construct of the developed rubric and the scores to be taken from each stage are presented below.

Table 1. Problem solving rubric

| Point | Understanding the problem, determining its limitations and difficulties. |
| :---: | :---: |
| 4 | Can thoroughly define the construct and limitations of the problem and can show the ambiguous limitations. |
| 3 | Can thoroughly define the main construct and limitations of the problem. |
| 2 | Cannot define some aspects and limitations of the problem. |
| 1 | Cannot define many aspects and limitations of the problem. |
|  | Presenting applicable and consistent ways of solution for the problem. |
| 4 | Can define creative and usable ways of solution for the problem under discussion. Solutions target the critical points. |
| 3 | Can present different ways of solution including the main critical points of the problem. |
| 2 | Can present ways of solution including some critical points of the problem. |
| 1 | Can present ways of solution not including the critical points of the problem. |
|  | Selecting the ways of solution and making an adequate number of trials. |
| 4 | Can try the alternative ways of solution selected effectively and in a detailed manner. |
| 3 | Can establish a sound connection between their trails to solve problems and their problem solving knowledge. |
| 2 | Can prove that the alternative solutions they have selected are correct and useful. |
| 1 | Try ways of solutions not leading to conclusions; either never use some elements of a problem or use them erroneously. Select unsuccessful ways of solution. |
|  | When different ways of solutions have been tried, explaining and discussing why these ways of solution have been selected and revealing which approach is more comprehensive and useful. |
| 4 | Can present a comprehensive and clear summary of the reasoning leading to the selection of secondary solutions. This summary includes decisions showing the alternatives to be selected and which alternatives can be accepted as solutions. |
| 3 | Can define the structure of secondary solutions. This definition is the ultimate selection apparently revealing the structure of different solutions. |
| 2 | Can define the structure of secondary solutions. Yet, this definition does not reveal the actual structure of the secondary solutions. |
| 1 | Cannot define reasonable ways of determining the importance of different solutions relative to each other. Cannot offer a reasonable explanation showing the missing or strong sides of tried and given up alternative solutions. |

As can be seen, the student is expected to revise understand the problem and to recognize limitations, explicit and implicit sides and difficulties of it in the first stage. This is related to the extent to which the student has internalized prior learnings related to the topic or the concept.
In the second stage, the student is expected to produce creative and useful solutions targeting critical points to solve the problem. It requires the construction of a mathematical model as well as reasoning because the student; by constructing a model, emphasizes the critical points of the problem. The approaches adopted by him/her while establishing connections can be evidence of the strategies he/she has used. Thus, the student's success in the second stage depends on his/her skill of reasoning and developing strategies.

In the third stage, the student is asked to try the alternative solutions he/she has selected in an effective, valid and detailed manner. $\mathrm{He} /$ she is expected to evaluate whether the results he/she has achieved are in compliance with the conditions.
In the fourth stage, individuality comes to the fore. Here, the student is expected to discuss the solutions and to explain why he/she has selected this solution or these solutions. In addition to this, he/she is asked to make inferences from the results he/she has reached. $\mathrm{He} /$ she is expected to make efforts to expand or develop the problem on the basis of his/her inferences. In this way, the ways of relating the solution to the real life or other areas can be opened. In short, he is at least expected to seek answers for such questions as "If it were" to expand the problem.
All the preparation and application works conducted with the students were led by the first author. Delivery of the math lessons in both of the classes by the researcher facilitated organization of the classes.
Before the application, first some information about the experimental and control group students' level determination exam results taken by them in $6^{\text {th }}, 7^{\text {th }}$ and $8^{\text {th }}$ grades was collected. In this regard, the scores taken by the students and the number of their correct answers in math sections were determined. In addition to this, the students' grade point averages for the math course at the end of the $8^{\text {th }}$ grade were tabulated. The school administration confirmed that the classes were formed in such a way as to include students from every level. As the application was conducted in an Anatolian high school, the students' high school entrance scores were close to each other; thus, not much difference between their levels was expected. In the second stage, in order to check their prior knowledge and to determine shortcomings, all the students in the control group and experimental group were administered a level determination exam (DBS). The questions of this exam were prudently prepared considering the questions involved in the "level determination exams" prepared by the Ministry of National Education for $7^{\text {th }}$ and $8^{\text {th }}$ graders, by making use of different resources in such a way as to capture all the learning areas in the elementary school $6^{\text {th }}, 7^{\text {th }}$ and $8^{\text {th }}$ grade curriculums (MEB, 2009). The questions were submitted to expert review and in line with the feedback given by the experts, some questions were discarded and some were corrected; thus, a test of 50 questions was formed. The developed DBS was piloted on 150 students in 2009-2010 school year. The data obtained from this administration were used to test the internal consistency of the test and some questions with low reliability were excluded from the test. Thus, the test consisting of 23 questions with high reliability was formed. The internal consistency coefficient of the test was found to be 0.822 . As a result, the test was considered to be highly reliable (Özdamar, 1999; Tavşancıl, 2006). The actual application of the test was performed in 2010-2011 school year with the 23 -item test and as a result of this application, two more questions testing the units of probability and statistics were discarded; thus, the 21 -item test was obtained. Moreover, the test questions were turned into classical exam questions so that better insights could be gained about the students' knowledge.
After the administration of DBS, revision of the subjects was made to help students overcome their difficulties and the critical points of the basic concepts were emphasized. After it was determined that the two classes were close to each other, it was decided that the two classes were comparable.
As it was assumed that the students had not utilized the problem solving process before, in the preparation stage, the control and experimental group students were presented with problem solving samples for four class hours. Here the main objective was to make the students familiar with the kinds of problems they would encounter for the first time. During the application, the problem solving stages were examined and discussed with the students one by one. The students were explained what is expected of them, methods and approaches to be followed in the solution and the strategies that could be used. While the problems solved in the class were constructed, a great care was taken for them to test the solution of verbal, algebraic and geometric questions. After the solution of the sample problems, the worksheets were handed out to the students including different problems and they were asked to solve these problems as homework.
Before the application, the experimental group students were divided into groups of six considering their SBS6, SBS7 and SBS8 scores, the number of correct math answers in these exams and math grade point averages. A great care was taken to include an equal number of male and female students in each group. At the beginning of the application, the classroom was arranged in $U$ form. The students were informed about what working in cooperation means, how they should work in cooperation and why the classroom was arranged in $U$ form.
After the arrangement of the classroom environment and before the application, in order to determine the students' performance, the Performance Measurement Scale 1 (PÖÖ1) was administered to both groups of students; thus, their performance before the application was determined.
Throughout the application, all the activities were performed with the aim of improving the students' performance. Both in the control and experimental groups, all the lessons were taught by the researcher. At the
end of the application process, all the students were administered the Performance Measurement Scale 2 (PÖÖ2) to determine the extent of the difference created by the application.
In the learning environment designed to conduct the application properly, a great emphasis was put on both within-groups and whole-class discussions in the experimental group and in the control group throughout the application process. While the activities in the experimental group were designed to enhance performance, the activities used in the control group were designed in relation to daily life, prior learning and other areas of learning. By means of assignments, projects and worksheets, the students were encouraged to interact with each other and through technologic and written resources; the students were encouraged to interact with specialists. In short, attempts were made to convert the classroom environment bordered with four walls to an environment connected with the state, science and technology and where the Web fountain is always turned on (Pat, 2000). In the learning environment designed for the experimental group, what was aimed was to make the students proceed from presentation to relating, from speaking to writing, from textbook to the generation of standard concepts, from recalling to doing, from reciting to understanding, from superficial evaluation to performance evaluation and from school test achievement to real life achievement (Alkan,2008).
The pilot study of the application was conducted in 2009-2010 school year and the actual application was conducted in 2010-2011 school year. The pilot study and actual application lasted for a total of 27 weeks. In high school, there are 4 class hours allocated to math course; thus, the study took 108 class hours. The study was kept limited to the learning areas of clusters, Cartesian multiplication, relations and functions in the $9^{\text {th }}$ grade curriculum.

## 3. Results

First the experimental and control groups were compared prior to the application. Whether the distribution of the mean scores taken from PÖÖ1 by the students met the normality hypothesis was examined and the results presented in Table 1 were obtained.

Table 2. Kolgomorov-Smirnov normality analysis results for the distribution of the mean scores taken from PÖÖ1 by the experimental and control group students

| Groups | N | Mean | Standard Deviation | Kolmogorov-Smirnov | p |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Experimental | 29 | 3.72 | 1.67 | 0.125 | 0.200 |
| Control | 29 | 3.87 | 1.92 | 0.136 | 0.182 |

$\mathrm{p}>0.05$.

As can be seen in Table 2, p value found as a result of the Kolmogorov-Smirnov test conducted for the experimental and control groups is higher than 0.05 . That is, PÖÖ1 mean scores belonging to the experimental and control groups showed a normal distribution. Then, in order to compare the mean scores of the experimental and control groups, $t$-test was run and the results are presented in Table 3.

Table 3. Results of the t -test run to compare the PÖÖ1 mean scores of the experimental and control groups

| Groups | N | Mean | Standard Deviation | t | p |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Experimental | 29 | 3.72 | 1.67 | 0.312 | 0.756 |
| Control | 29 | 3.87 | 1.92 |  |  |

As can be seen in Table 3, there is no statistically significant difference between the mean scores taken from the PÖÖ1 test by the experimental and control groups. Moreover, the mean scores are very close to each other.
After the application, in order to compare the performance levels of the experimental and control groups, first it was investigated whether the distribution of the mean scores taken from PÖÖ2 meets the normality hypothesis and it was found that it has a normal distribution.

Then, $t$-test was run to see whether the difference between the mean scores of the experimental and control groups is statistically significant and the results are presented in Table 4.

Table 4. Results of the t-test run to compare the PÖÖ2 mean scores of the experimental and control groups

| Groups | N | Mean | Standard Deviation | t | P |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Experimental | 29 | 6.31 | 2.18 | 5.959 | 0.000 |
| Control | 29 | 3.39 | 1.48 |  |  |

As can be seen in Table 4, there is a statistically significant difference between the PÖÖ 2 mean scores of the experimental and control groups. This difference is in favor of the experimental group. Thus, it can be argued that teachers' supporting conceptual thinking and attaching importance to students' interpretations as well as math course learning activities have positive contributions to the development of performance. With such applications, students' problem solving achievement and conceptual understanding achievement can be increased (Kazemi, 1998).
After the application, comparison of the pretest and posttest mean scores of the experimental group students was also made. Their mean scores taken from PÖÖ1 and PÖÖ2 exhibited normal distributions and t-test was run to determine whether there is a significant difference between the two mean scores and the results are presented in Table 5.

Table 5. Results of the t-test run to compare the experimental group students' PÖÖ1 and PÖÖ2 mean scores

| Groups | N | Mean | Standard Deviation | t | P |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Experimental1 | 29 | 3.72 | 1.67 | 5.075 | 0.000 |
| Experimental2 | 29 | 6.31 | 2.18 |  |  |

As can be seen in the Table the level of significance is $\mathrm{p}<0.05$. That is, there is a significant difference between the pretest and posttest mean scores of the experimental group students in favor of the posttest mean score. Thus, it can be argued that the methods and approaches used in the application led to a significant difference.
It was also investigated from which sub-components of performance this development resulted and which sub-components can be connected with this development. For this purpose, the experimental group students' PÖÖ1 and PÖÖ2 mean scores belonging to each component were compared. To this end, Wilcoxon Signed Rank test, one of the non-parametric tests, was used and the results are presented in Table 6.

Table 6. Wilcoxon signed rank test results in relation to the performance components of the experimental group students' PÖÖ1 and PÖÖ2 mean scores

| Component | Posttest-Pretest | N | Mean Rank | Rank Sum | z | p |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Component 1 | Negative Rank | 0 | 0 | 0 | 4.706* | 0.000** |
| (Understanding) | Positive Rank | 29 | 15.00 | 435 |  |  |
|  | Equal | 0 |  |  |  |  |
| Component 2 | Negative Rank | 9 | 10.72 | 96.5 | 1.778* | 0.075 |
| (Establishing | Positive Rank | 16 | 14.28 | 228.5 |  |  |
| connections) | Equal | 4 |  |  |  |  |
| Component 3 | Negative Rank | 8 | 9.19 | 73.5 | 2.951* | 0.003** |
| (Application) | Positive Rank | 20 | 16.63 | 332.5 |  |  |
|  | Equal | 1 |  |  |  |  |
| Component 4 | Negative Rank | 3 | 9.17 | 27.50 | 3.884* | 0.000** |
| (Inferences and | Positive Rank | 24 | 14.60 | 350.50 |  |  |
| development) | Equal | 2 |  |  |  |  |

The analysis results presented in the Table show that there are statistically significant differences between the pretest and posttest mean scores for the $1^{\text {st }}, 3^{\text {rd }}$ and $4^{\text {th }}$ components of the PÖÖ 1 and PÖÖ 2 performance measurement tools. When the mean rank and rank sum are considered, it is seen that these differences are in favor of PÖÖ2. On the other hand, there is no significant difference between the pretest and posttest mean scores for the $2^{\text {nd }}$ component. That is, no significant difference was observed in relation to the component of establishing connections. This might be because the students were familiar with mathematical model connections.

The mean scores obtained for the sub-components of PÖÖ 2 increased when compared to the sub-components of PÖÖ1. The reason for this increase might be problem solving activities and interpretations of these solutions.

## 4. Summary and Conclusion

It is seen that at schools in many countries and in international exams, the performance-based evaluation has come to the fore. In these countries, a great deal of research on performance evaluation and development has been conducted. However, in our country, there is not much research focusing on the performance-based evaluation. In our country, what is understood from the evaluation of performance is end-of-term assignments and long-term assignments. In fact, many of these assignments can be easily done by cutting and pasting operation. However, performance assignments should encourage students to understand the conditions of scientific work and to think and reason during the problem solving process (Solano-Flores \& Shavelson, 1997; Palm, 2008). Thus, it can be maintained that there is a widespread misconception about the concept of performance and performance evaluation. The first thing to be done by the Ministry of National Education desiring to initiate the performance-based evaluation system should be to correct this misconception because correction of such misconceptions seems to be more important than the initiation of the application process.
In today's education system, the concept of learning environment has gained a new meaning. In this regard, it is emphasized that new approaches should be adopted for the arrangement of classrooms and classroom equipment and that the classroom environment should be just a part of this new learning environment. Seen from this perspective, the learning environment is considered to be an environment where students can have access to all information, can express what they think comfortably and which helps the expression of individual differences. The learning environment should also be adjusted according to preferred learning approaches, preferred ways and techniques of learning and evaluation methods to be used. Thus, when a learning environment is designed to enhance students' performance and learning activities are developed in line with this new environment, it becomes clear that more in-depth teaching of each subject is necessary (Darling-Hammond \& Laura, 2008). For example, the results of the Program for International Student Assessment (PISA, 2006) revealed that the United States of America took the $35^{\text {th }}$ place out of 40 countries. This negative result motivated authorities to explore the curriculums of more successful countries. As a result of the analysis of the collected data, it has been found that the curriculums of these more successful countries not only focus on teaching of subjects but also on more in-depth analysis of subjects, on the development of students' reasoning skills and on the provision of practice opportunities (Darling-Hammond \& Laura, 2008). As can be seen from these findings, what is more important in PISA is performance-related variables. Accordingly, if international success is to be achieved, then it seems to be inevitable to design a learning environment, to select methods and techniques and assessment tools that will be conducive to the performance development of students.
The measurement tools to be used to evaluate performance need to measure the student's knowledge, the skill of using methods and techniques and reasoning skill. In order to be successful, education systems need to take these variables into consideration. On the other hand, it is very difficult and even impossible to measure these variables by using multiple-choice or completion questions or by using tests only based on operations and recall. What needs to be done is to develop and apply measurement tools that can measure these variables. Particularly in mathematics, in order to be able to measure students' performance in each of these variables, it is necessary at least to determine their skills and reasoning ability while dealing with non routine and open-ended questions. Moreover, as a result of such measurement, it can be possible to gain insights into students' many skills such as modeling, reasoning, making use of their knowledge and developing their knowledge that are also very useful in the real life. In other words, such assessment approaches also prepare students for the real life.
For a soft transition to new measurement approaches in today's world within the context of a system change, at the outset of this change, while developing measurement tools, a few short open-ended questions and non routine questions can be included so that the behaviors exhibited by students while solving such problem can be analyzed to shed some light for the further stages. Yet, before this, students' skill of producing opinions should be considered important and it should be evaluated. In addition to this, short-time and long-time term papers can be designed in such a way as to measure the above-mentioned variables. In this connection, curriculums can be rearranged.
Aside from the pilot study, the application lasted for one school year. This time period seems to be not very long as the students need to adapt a new approach. Yet, the results obtained at the end are more positive than the expected. If loner experimental studies can be conducted, more positive results in relation to student performance and achievement can be obtained.

## References

Alkan, H., \& ve Ceylan, A. (2008). Learning environment and program design for mathematical thinking development of mathematics teacher candidates. D.E.Ü. Buca Education Faculty. DPT Project No: 2003 K120360.
Baker, E. L. (1997). Model-based performance assessment. Theory into Practice, 36(4): New Directions in Student Assessment, 247-254. https://doi.org/10.1080/00405849709543775
Baxter, J. (2001). Enhancing student achievement on performance assessments in mathematics. Learning Disability Quarterly, Jan 1.
Brecher, N. (2007). Performance enhancement: Treat the cause of poor work, not just the symptoms (good to great). Journal of Property Management, July-August, 2007.
Darling-Hammond, L., \& McCloskey, L. (2008). Assessment for learning around the world: What would it mean to be internationally competitive? Phi Delta Kappan, 90(4), s.263-272. https://doi.org/10.1177/003172170809000407
Elliott, S. N., \& Fuchs, L. S. (1997). The utility of curriculum based measurement and performance assessment as alternatives to. School Psychology Review, 26(2).
Holtz, G., \& Malen, M. L. (2007). Open-ended math problems. Retrieved from http://www.fi.edu/school/ math2/index.html
House, J. D., Keely, E. J., \& Hurst, R. S. (1996). Relationship between learner attitudes, prior achievement, and performance in a general education course: A multi institutional study. International Journal of Instructional Media, 23(3), 257-271.
Kazemi, E. (1998). Discourse that promotes conceptual understanding research into practice. Teaching Children Mathematics, March, 410-414.
Kılıç, S. D. (2011). Researching assessment and improvement approachment of students' performance level that exhibited in mathematics courses. Unpublished doctoral thesis. Dokuz Eylül University, Institute of Educational Sciences, İzmir.

Kitchen, R., Cherrington, A., Gates, J., Hitchings, J., Majka, M., Merk, M., \& Trubow, G. (2002). Supporting reform through performance assessment. Mathematics Teaching in the Middle School, 8(1), September.
Lianghuo, F. (2011). Performance assessment in mathematics: Concepts, methods, and examples from research and practice in Singapore classrooms. Singapore: Pearson Education South Asia.
M.E.B. (2009). Math course 6-8. Class Curriculum. Ankara: M.E.B. Publications.
M.E.B. (2009). PISA international student achievement evaluation program. Retrieved from http://earged.meb.gov.tr /pisa/dil/ tr/ sunum.html
Marzano, R. J., Pickering, D. J., \& MCTighe, J. (1993). Assessing student outcomes. Library of Congress Cataloging - in - Publication Data. USA: ASCD.
Niemi, D. (1997). Cognitive science, expert-novice research, and performance assessment. Theory into Practice, 36(4): New Directions in Student Assessment, 239-246. https://doi.org/10.1080/00405849709543774
Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. Practical Assessment, Research \& Evaluation, 13(4).
Pat, H. (2001). The changing role of the teacher. The Journal, 26(November).
Pelfrey, R. (April, 2000). Open ended questions for mathematics. Service of the Arsi resource collaborative, University of Kentucky. Retrieved from http://www.uky.edu/OtherOrgs/ARSI/www.uky.edu/pub/arsi/openresponsequestions/geometyorq.pdf
Slater, T. (2007). Performance assesment. Retrieved from http://www.flaguide.org/ extra/download/ cat/ perfass/ perfass.pdf
Spring branch independent school district. (2010). Mathematics performance assessments. Retrieved from http://www.springbranchisd.com/instruc/math/assessment.htm
Sweet, D. (1993). Performance Assessment (Office of Research, Office of Educational Research and Improvement (OERI) of the U.S. Department of Education). Retrieved from Education Research Consumer Guide: http://www2.ed.gov/pubs/OR/ConsumerGuides/perfasse.html

Tavşancıl, E. (2006). Measuring attitudes and data analysis with SPSS. Ankara: Nobel Publishing.
Woodward, J., Monroe, K., \& Baxter, J. (2001). Enhancing student achievement on performance assessments in mathematics. Learning Disability Quarterly, 24(1), 33-46. https://doi.org/10.2307/1511294

## Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.
This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).

