Scaling of secondary school students' evaluation preferences according to learning approaches

Filiz İsnaç, *MEB Teacher,* Turkey, *filiz.temel30@gmail.com* ORCID: 0000-0002-2578-912X
Celal Deha Doğan, *Ankara University,* Turkey, *dehadogan@gmail.com* ORCID: 0000-0003-0683-1334

Abstract. The aim of this study is to determine the evaluation preferences of secondary school students adopting deep and superficial learning approach by using pairwise comparison method. The study was carried out with 444 students chosen from 9th, 10th, 11th and 12th grade students in four different types of secondary education in Mersin. The case V and case III models of the scaling method based on Thurstone's (1927) Law of Comparative Judgment, which can be applied in cases where the items can be given to participants in pairs, are used to answer the research questions. As a result of the research, it was determined that students adopting both deep and superficial learning approaches prefer the questions measuring the skills at remembering level with higher chance success such as matching, true or false and multiple choice in the first place, performance tasks and long answer questions in the last place while determining their academic success. When this result was examined according to the overall and grade levels of secondary students, it was observed that it did not change except for minor differences.

Keywords: Evaluation preferences, learning approaches, scaling

INTRODUCTION

The quality of the teaching and learning process depends on a number of factors, such as the characteristics of the teacher and the student, as well as factors related to the learning environment. The teacher's knowledge and skills, the teaching methods and techniques used, the learning styles of the student, the strategies and learning approaches used while learning, and the assessment methods they use are the main variables that influence the student's academic achievement. According to Kuzgun and Deryakulu (2004), it is an important element that teachers who plan and organize an educational environment and monitor learning outcomes have sufficient knowledge of these variables that reveal individual differences and use them in a way that increases the learning capacity of students.

Assessment activities in the classroom environment create differences on the students' learning and their success accordingly. Philips' (1999) study supports this judgment. According to the findings of this research; the students choosing tasks to solve complex problems were more successful than those choosing problems that were easier and faster to solve. This situation indicates that the concept of evaluation preference should be considered as a variable affecting the success of the teaching process. The assessment tools used to determine the academic achievement of students, the type and level of questions, the stage of scores, in short, evaluation methods and ideas about the characteristics of those methods are generally called evaluation preferences (Birenbaum, 1994; 1997).

Taking into account individual differences, while structuring the teaching process, is important in terms of increasing the quality of teaching. Characteristics of learning, such as learning approaches, are also individual differences that affect student achievement. Recent studies have shown that not all students learn the same way and follow different approaches. While some of the students make a lot of effort to grasp the subject, some of them just try hard enough to pass the course (Biggs, 1999; cited in Jones, 1999). Marton and Saljo (1976a, 1976b) first introduced the concepts of deep and superficial learning approaches with their studies on

how students attempt to comprehend a reading text given to them and how they handle learning. Deep learning approach is defined as the tendency to be eager to learn, to concentrate on the content of the subject and to have a critical interaction with the subject, to establish a relationship between their previous learning and their new learning, to internalize the concepts by linking them with their daily lives, to establish a cause-effect relationship between events and to examine the logic of the subject. On the other hand, the superficial learning approach is defined as the desire to do only the requirements of the given job, memorization of the possible issues in the exam, the inability to distinguish between principles and examples, and the tendency to see learning as an external imposition (Byrne, Flood and Willis, 2002).

It is possible to state that students' preferences regarding teaching processes and assessment methods affect learning perceptions and approaches that are effective in their success (Biggs, 2003; Brenbaum, 2003; Struyven et al., 2005). In other words, how students handle learning and how the environment influences them has a decisive effect on learning products and assessment preferences (Ekinci, 2008). The impact of learning approaches and assessment preferences on the teaching and learning process has spread to research on the subject. Research on this subject shows that learning approaches make a significant difference on students' academic achievement and evaluation preferences (Gijbels and Dochy, 2006; Mayya, Rao and Ramnarayan, 2004; McManus, Richards and Winder 1999; Sambell et al., 1997; Scouller, 2000).

Determining and taking into account the variables affecting student achievement while structuring the teaching process is important in terms of improving the quality of teaching. It is therefore important to determine the evaluation preferences that have a significant impact on the achievement of students and to reveal differences in their preferences. In this regard, research findings are expected to draw attention to the use of evaluation preferences in the organization of curriculum, planning of teaching process and activities.

Considering that assessment activities in classroom environment affect and guide learning approaches, it is expected that the research findings would guide teachers about which assessment methods they should use during the preparation of assessment activities. In addition, it is thought that the results of the research will contribute to the ongoing discussions in the field about the types of questions that should be used in the examinations.

When the literature was analyzed, a limited number of studies on evaluation preferences were found. It is seen that these studies are mostly either evaluation preferences of the teachers and prospective teachers (Bal, 2012a; Bal, 2012b; Büyüköztürk and Gülbahar, 2010; Doğan, 2011; Gelbal and Kelecioğlu, 2007) or correlative studies aimed at determining the relationship between the characteristics related to learning and evaluation preferences (Bal Doğan, Atmaca and Aslan, 2013; Doğan and Kutlu, 2010). The information provided in the correlation studies showing the relationship between the variables is extremely important. However, these types of studies on evaluation preferences reveal the existence and level of a significant relationship between the variables, rather than providing the necessary information about which exam types are preferred and their rankings. Scaling studies contribute significantly to the field in order to address this issue. The lack of a scaling study to determine the evaluation preferences of secondary school students reveals the need for this issue. In this regard, it is thought that the study will contribute to addressing this deficiency in the field, and this situation is seen as another point that increases the importance of the study.

The aim of this study is to determine the evaluation preferences of secondary school students who take a deep and superficial approach to learning using a pairwise comparison method. In line with this general aim, first of all, the evaluation preferences of secondary school students were determined in general and then separately according to their grade levels (9th, 10th, 11th and 12th grade).

METHOD

Research Model

This study was designed as survey model. Survey models are research approaches aiming at describing a past or present situation as it is (Karasar, 2014).

Study Group

The study was carried out in a total of four state high schools in Mersin, an Anatolian high school, a science high school, an imam hatip high school and a vocational high school on 465 volunteer students from 9th, 10th, 11th and 12th grade. When the data set was analyzed, it was seen that 21 students scored equal from the Study Process Questionnaire (R-SPQ-2F). Since there was a strong negative relationship between the two subscales of the Evaluation Preference Pairwise Comparison, it was suspected that these students did not respond carefully to the questions. For this reason, the responses of the students to the Evaluation Preference Pairwise Comparison (PC) form were examined and the internal consistency was calculated. This value, known as the Kendall's zeta statistics (1948), is used to calculate the internal consistency of the judgments of the observers and to assess their significance. The chi-square value of the response of 21 participants was calculated using this method. Each of the calculated values was compared with the value of 21 degrees of freedom in the chi-square table. Since the values obtained are greater than the value determined from the table ($X_{21}^2 = 32.671$) with a reliability level of 0.95, it was found that the participants were not consistent at reliability level of 0.95 when answering the form. It was therefore decided to exclude the participants from the study group. Analyzes were carried out on the remaining data set, that is, the data of 444 participants. In the determination of the study group, maximum variation sampling method, which is one of the purposive sampling methods, was used. As a study area, different grade levels and school types were identified within the scope of the study.

Data and Collection

Two measurement tools were used for the stated purpose in the process of collecting the research data. The first tool is the Study Process Questionnaire (R-SPQ-2F), developed by Biggs (1987) and revised by Kember, Biggs and Leung (2004) and adapted to Turkish culture by Çolak and Fer (2007). The other one is the Evaluation Preference Pairwise Comparison Form, which was developed by the researcher to determine the evaluation preferences of the students. Firstly, a comprehensive literature review was conducted for the Evaluation Preference Pairwise Comparison Form developed by the researcher. In addition, the opinions of the teachers have been used to determine the methods that the students have experience of and to exclude the methods that they are not familiar with. In this respect, 8 assessment methods (multiple choice questions, short answer questions, long answer questions, true or false questions, matching questions, gap filling questions, performance tasks (research-based reports), performance tasks (research-based presentations)), which are widely used in schools and in which students have experience, have been determined. The form was then finalized by consulting 2 assessment and evaluation experts. After the form was prepared, 25 students were pretested and the final form was created by a revision based on the deficiencies. Practices have been carried out by the researcher herself. Necessary explanations have been provided in order to avoid confusion as to the methods established prior to the process.

Data Analysis

To find answers to research questions, analyzes were carried out using case V and case III model of the scaling method through pairwise comparisons based on Thurstone's (1927) Law of Comparative Judgment. This method, which can be applied in any case where stimuli can be given to participants in pairs, is preferred for scaling many features in the affective domain (Turgut and

Baykul, 1992). In this scaling method, when scaling with case V model, the frequency Matrix (F) is created by first calculating the frequencies of the participants' responses to pairwise comparisons. In the next step, these values are divided by the sum of the number of participants to form the proportions matrix (P). In the second step, the unit normal deviations matrix (Z) is generated by finding the standard values (z) corresponding to each value in the proportions matrix. In the third step, the means of the standard values (z) are calculated and the scale values are found. In the last step, the absolute value of the smallest of the scale values is added to all values. In this way, these values are sorted by shifting the starting point of the scale values to zero.

The scale values obtained as a result of the calculations made according to the case V model of the pairwise comparison method are based on some assumptions along with the participants' judgments. The internal consistency of scaling is checked to analyze the consistency of participants' judgments and the availability of assumptions. For this purpose, the mean of the difference between the observed ratio and the theoretical ratio is calculated, namely the mean error. If the calculated mean error is small and close to zero, it is interpreted that the internal consistency of the scale values is high, that is to say, the scale values are not reliable. In other words, the mean error value is a measure that shows an agreement between the observed value and the empirical value. Chi-square test is used to check the significance of the mean error. According to Guilford (1954), the case III model should be used if the value calculated by the chi-square test is significant. The steps that are followed when scaling with this method are as follows:

1. Standard deviations of the column elements in the matrix Z are found with the following formula;

$$V_j = \frac{1}{K} \sqrt{K \sum_{k=1}^{K} z_{kj}^2 - \left(\sum_{k=1}^{K} z_{kj} \right)^2}$$

2. KC constant is found with the following formula;

$$KC = \frac{2K}{\sum_{j=1}^{K} \frac{1}{KV_j}}$$

3. The standard deviations of discrimination judgments are calculated by the following formula;

$$\sigma_j = \frac{K.C}{K.V_j} - 1$$

4. Variance values are calculated by squaring the standard deviation. The sum of the variance matrix is formed by taking the pairwise sum of these numbers. Then, the square root matrix of the sum of variance is formed by taking the square roots of the values found in this matrix.
5. S matrix is created by multiplying sum of variance with each element on the main diagonal of the square root matrix and elements located on the principal diagonal of the unit normal deviations matrix.
6. Column averages of the elements in the S matrix are taken and the scaling is completed by shifting the starting point of the smallest mean to zero.

According to information provided, eight assessment methods were given in pairwise form to each student included in the study, and they were asked to mark their most preferred one by comparing two options. Then, the frequency of the pairwise comparisons of the participants was calculated and the frequency matrix was obtained. Eight evaluation methods were coded by matching the letters [A-H] to be useful for constructing the frequency matrix. Multiple choice questions (A), short answer questions (B), long answer questions (C), true or false questions (D),

matching questions (E), gap filling questions (F), performance tasks (research-based reports) (G) and performance tasks (research-based presentations) (H) are expressed with letters. The values found in the frequency matrix were divided by the total number of people, and the proportions matrix was obtained. The standard values (z) corresponding to each element in the proportions matrix were calculated in the next step, and the unit normal deviation matrix was formed. The scale values of this matrix were calculated by finding column averages. In the last step, the absolute value of the smallest of the scale values obtained was added to all values and the scale values were determined according to the case V model. After these steps, the mean error amount was calculated to check whether or not the assumptions of the case V model were met and to determine the consistency of the judgments of the participants. The significance of this error was checked by the chi-square test. When the amount of error was significant, scaling was done with case III model.

The first question of the study was asked as "What are the scale values of the evaluation preferences of secondary school students according to the pairwise comparison method?" and as a result of the analyses, the amount of error was determined to be significant, and the analyzes were continued using the case III model [$\chi^2_{21} = 57.127$, p<.05]. Similarly, it was found that the amount of error was significant as a result of the analyzes made for the 3rd sub-problem of the study, "What are the scale values of the evaluation preferences of secondary school students adopting the superficial learning approach according to the pairwise comparison method?", and the analyzes were continued using the case III model [$\chi^2_{21} = 38.470$, p<.05]. According to the Case III model, standard deviations of discriminatory judgments were determined for these two research questions. Variance values were calculated by finding the squares of the standard deviations determined. The sum of variance matrix was obtained by taking the pairwise sum of the variances found. The square root of the sum of variance was obtained by finding the square root of each element in the matrix obtained. Each element on the main diagonal of this matrix and the elements on the same diagonal of the unit normal deviation matrix were multiplied with each other, and S matrix was formed.

FINDINGS

The first question of this study is as "What are the scale values of the evaluation preferences of secondary school students according to the pairwise comparison method?" In order to answer this question, firstly, the case V model of the Thurstone's Law of Comparative Judgment was scaled. The mean error value was then calculated to test the case V model assumptions and to determine the degree of consistency of the participants' responses. According to the calculations, this value was found to be 0.014. The significance of this error was examined by chi-square test and found to be significant [$\chi^2_{21} = 57.127$, p<.05]. For this reason, analyzes were performed using case III model.

Scale values and rankings for the evaluation preferences of secondary school students obtained as a result of calculations made using the Case III model are shown in Table 1.

Table 1. *Scale values of evaluation preferences and stimulant orders*

Evaluation Preferences	Scale Values	Stimulant Order
Matching Questions (E)	0.994	1
True or False Questions (D)	0.903	2
Multiple Choice Questions (A)	0.638	3
Short Answer Questions (B)	0.583	4
Gap Filling Question (F)	0.457	5
Performance Tasks (Research-based reports) (G)	0.368	6
Performance Tasks (Research-based presentations) (H)	0.304	7
Long Answer Questions (C)	0.000	8

When Table 1 is examined, it is seen that the questions with matching type have the highest scale value. This shows that secondary school students prefer to be evaluated with matching questions in determining their academic success. True or false questions are noteworthy as a preference that follows matching questions with a close scale value. These choices are followed by multiple choice questions, short answer questions, gap filling questions, performance tasks (research-based reports), performance tasks (research-based presentations) and lastly long answer questions.

The second question of this research is as "What are the scale values of secondary school students' evaluation preferences by paired comparison method according to grade level? ". In order to answer this question, firstly, the case V model of Thurstone's Law of Comparative Judgment was scaled for each grade level. The mean error amounts were then found to test the assumptions of the case V model and to assess the degree of consistency of the participants' responses. The significance of these error amounts was examined by chi-square test, and it was determined that they were not significant. In other words, the internal consistency of the scale values has been determined to be at the desired level. For this reason, scaling based on the case III model was not required.

Scale values and rankings related to evaluation preferences of secondary school students based on grade levels, obtained as a result of calculations made using Case V model, are given in Table 2, Table 3, Table 4 and Table 5.

Table 2. *9th grade students' scale values of evaluation preferences and stimulant orders*

Evaluation Preferences	Scale Values	Stimulant Order
Matching Questions (E)	0.729	1
True or False Questions (D)	0.609	2
Gap Filling Question (F)	0.386	3
Matching Questions (E)	0.329	4
Performance Tasks (Research-based reports) (G)	0.275	5
Short Answer Questions (B)	0.266	6
Performance Tasks (Research-based presentations) (H)	0.214	7
Long Answer Questions (C)	0.000	8

When Table 2 is examined, it is seen that matching type questions have the highest scale value. This shows that secondary school students prefer to be evaluated with matching questions in determining their academic success. True or false questions are again noteworthy as a preference that follows matching questions. These choices are followed by gap filling questions, multiple choice questions, performance tasks (research-based reports), short answer questions, performance tasks (research-based presentations) and lastly long answer questions.

Table 3. *10th grade students' scale values of evaluation preferences and stimulant orders*

Evaluation Preferences	Scale Values	Stimulant Order
Matching Questions (E)	0.705	1
True or False Questions (D)	0.596	2
Short Answer Questions (B)	0.500	3
Multiple Choice Questions (A)	0.382	4
Gap Filling Question (F)	0.335	5
Performance Tasks (Research-based reports) (G)	0.215	6
Performance Tasks (Research-based presentations) (H)	0.148	7
Long Answer Questions (C)	0.000	8

When Table 3 is examined, it is seen that matching type questions have the highest scale value. This shows that secondary school students prefer to be evaluated with matching questions in determining their academic success. True or false questions are again preferred following matching questions. These choices are followed by short answer questions, multiple choice questions, gap filling questions, performance tasks (research-based reports), performance tasks (research-based presentations) and lastly long answer questions.

Table 4. *11th grade students' scale values of evaluation preferences and stimulant orders*

Evaluation Preferences	Scale Values	Stimulant Order
True or False Questions (D)	0.468	1
Matching Questions (E)	0.415	2
Multiple Choice Questions (A)	0.389	3
Short Answer Questions (B)	0.287	4
Performance Tasks (Research-based presentations) (H)	0.239	5
Performance Tasks (Research-based reports) (G)	0.206	6
Gap Filling Question (F)	0.108	7
Long Answer Questions (C)	0.000	8

When Table 4 is examined, it is seen that true or false questions have the highest scale value. This shows that secondary school students prefer to be evaluated with true or false questions in determining their academic success. Matching questions stand out as a choice following true or false questions. These choices are followed by multiple choice questions, short answer questions, performance tasks (research-based presentations), performance tasks (research-based reports), gap filling questions, and lastly long answer questions.

Table 5. *12th grade students' scale values of evaluation preferences and stimulant orders*

Evaluation Preferences	Scale Values	Stimulant Order
Multiple Choice Questions (A)	0.949	1
True or False Questions (D)	0.886	2
Matching Questions (E)	0.860	3
Short Answer Questions (B)	0.777	4
Gap Filling Question (F)	0.535	5
Performance Tasks (Research-based reports) (G)	0.483	6
Performance Tasks (Research-based presentations) (H)	0.350	7
Long Answer Questions (C)	0.000	8

When Table 5 is examined, it is seen that multiple choice questions have the highest scale value. This shows that secondary school students prefer to be evaluated with multiple choice questions in determining their academic success. True or false questions are the choice following the multiple choice questions. These preferences are followed by short answer questions, gap filling questions, performance tasks (research-based reports), performance tasks (research-based presentations) and lastly long answer questions.

When the findings of the second research question are analyzed, the evaluation methods preferred by secondary school students to determine their academic success are parallel to their grade level. Scale values were generally similar to each other with few exceptions. The differences may also be due to the size of the study group. Additionally, the first two rankings of 9th, 10th, and 11th grade students were matching and true or false question types in the evaluation preferences, while it is noteworthy that multiple choice questions take the first place among the 12th grade students. It can be said that this is an expected situation considering that 12th grade students are preparing for university entrance exams since university entrance exams consist of multiple choice questions. As a result, students are constantly familiar with these types of multiple-choice questions while getting prepared for the exam. On the other hand, it can be said that 9th, 10th and 11th grade students are directed towards true or false and matching questions which have high academic success and chance success rates. It is noteworthy that long answer questions are the last choice at all grade levels. The reasons for this may be long answer questions based on interpretation, being away from memorization, taking more time to solve compared to other questions, not having the chance to solve the question by chance success and anxiety of objective scoring. In addition, the fact that performance tasks are preferred before long answer questions at all grade levels is another important finding. The reason for this may be that performance tasks are not being used properly in schools, and that students think they can get their assignments ready from the Internet or have others do it.

The third question in this study is, "What are the values of the assessment methods that they prefer when determining the academic achievement of secondary school students according to the learning approach adopted by the pairwise comparison method?". To answer this question, students are divided into two groups, primarily those who adopt a deep and superficial learning approach based on data from the Study Process Questionnaire (R-SPQ-2F). For both groups, the data obtained from the Study Process Questionnaire (R-SPQ-2F) were scaled with the case V model of Thurstone's Law of Comparative Judgment. The mean error amount was then calculated to test the assumptions of the case V model and to check the degree of consistency of the participants' responses. The significance of this error was tested by chi-square test. In case of significant error, scaling was continued using case III model.

The assessment preferences of the students who adopted the deep learning approach were scaled with the case V model according to the pairwise comparison method. The analyzes were carried out in a similar way on 213 people who adopted the deep learning approach. The assesment preferences of the students adopting the superficial learning approach were scaled with the case III model according to the pairwise comparison method. The analyzes were conducted in a similar manner on 231 individuals adopting the superficial learning approach.

In Table 6, scale values and stimulant orders related to evaluation preferences of students adopting deep and superficial learning approach are presented in a single table in order to see the differences between them more clearly.

When Table 6 is examined, it is seen that there is not much difference between the evaluation preference rankings of the students adopting both deep and superficial learning approach. While the students adopting the deep learning approach preferred short answer questions in the 3rd place, the students adopting the superficial learning approach preferred the multiple choice questions in the same order. This can be attributed to chance success in multiple choice questions. In addition, students adopting a deep learning approach preferred presentation-based performance tasks as a priority over report-based performance tasks. The priority preference of students adopting a superficial learning approach to report-based performance tasks can be attributed to their tendency to choose the least effort and the fastest way. According to the students who take this approach, the presentation of the given assignment or prepared material in front of the class in detail can be seen as a waste of time and a difficulty.

Table 6. *Scale values and stimulant orders of evaluation preferences of students adopting deep and superficial learning approach*

Evaluation Preferences	Scale Values		Stimulant Order	
	Deep	Superficial	Deep	Superficial
Matching Questions (E)	0.444	1.263	1	1
True or False Questions (D)	0.370	1.228	2	2
Short Answer Questions (B)	0.327	0.735	3	4
Multiple Choice Questions (A)	0.324	0.841	4	3
Gap Filling Question (F)	0.208	0.617	5	5
Performance Tasks (Research-based presentations) (H)	0.173	0.318	6	7
Performance Tasks (Research-based reports) (G)	0.097	0.614	7	6
Long Answer Questions (C)	0	0	8	8

As a result, it can be said that for both groups, when the academic achievement of students is determined, they prefer to be evaluated with question types that measure their lower-order thinking skills.

DISCUSSION and CONCLUSIONS

In the findings of the first question of the study, the scale values of secondary school students' evaluation preferences were determined according to pairwise comparison method, and these preferences were listed.

According to the results of the research findings, students preferred matching and true or false questions mostly in the first two ranks, which measure remembering level functions with

high chance success rate. The chance success in such questions can be seen as a reason for preference. Again, the choice of multiple-choice questions where chance success has a significant effect reinforces this possibility. In addition, the fact that students find such questions simpler, more understandable and fair may have played a role in their preferences. The findings of a study on secondary school students showed that students prefer multiple-choice exams because they find it simple, interesting, attractive, understandable and fair (Zeidner, 1987). Anderson's (1987) study on university students also produced similar findings. In addition, research findings have shown that students prefer methods that require higher-order mental processes, such as performance tasks and long answer questions in determining their academic success. This result shows that secondary school students avoid methods that measure skills such as problem solving, interpretation, analytical thinking, researching, critical thinking, associating knowledge with new situations and using it in daily life, whereas they tend to use methods that measure behavior at the level of remembering knowledge. The reason for this situation may be that activities involving mental processes such as interpreting information, associating it with everyday life, conducting research are not included in the teaching process. Indeed, this interpretation is reinforced by the findings of Biggs (1999) and Gijbels and Dochy (2006). The lack of such assessment methods in the university placement exam may have led students to turn to other types of questions. In addition, the fact that students have to do research and study outside the school and have to present their study in front of the classroom can be seen as the reasons why they prefer to do less work. Doğan (2016), in his study to determine the evaluation preferences of university students, determined that students are more likely to choose multiple-choice and true or false questions. The findings of the study are consistent with the findings of the research in this regard.

It may therefore be beneficial for teachers to carry out an exam including different types of questions in the evaluation of students in order to contribute to an increase in student success. In the findings of the second question of the study, the scale values of the evaluation preferences of secondary education students were determined according to the grade level by using pairwise comparison method, and these preferences were listed. According to the findings, students in grades 9, 10 and 11 preferred matching and true or false questions in the first two ranks in determining their academic success, and this situation differed in 12th grade students. According to this, 12th grade students stated multiple-choice questions as the first choice, unlike other grade levels. At all class levels, long answer questions were determined as the last choice. One of the reasons for this is that students often face multiple-choice questions in large-scale examinations. Thus, in large-scale examinations held in Turkey, focusing on question structures that measure high-order thinking skills, in addition to multiple-choice question types, may contribute to the preference of question types that measure higher-order mental processes.

In the findings of the third question of the study, scale values of evaluation preferences were determined according to the learning approach adopted by secondary school students, and these preferences were listed. The research findings have revealed that students who adopt both a deep and superficial learning approach are similar in their preferred methods of determining their academic success. Although the scale values of the students' preferred methods differed, it is another result that their rankings did not generally change. While the students adopting deep learning approach are expected to prefer the methods that measure higher-order mental skills due to their nature, it is an important result obtained from the research that they tend to prefer methods measuring behaviors at the level of remembering, similar to students adopting superficial learning approach. This conclusion, which contradicts the literature, shows that there are serious problems in the education system in Turkey, and that even students with potential tend to superficial learning. In this context, it is recommended that the Ministry of National Education should take the necessary measures to organize and implement both curricula and syllabuses in order to improve students' higher-order thinking skills.

In this study, it was determined that secondary school students preferred the methods limited to measuring behaviors at remembering level in the first place and higher-order thinking skills such as interpretation, research and critical thinking in the last place while determining their academic success. The reasons for this situation include many factors such as studying

strategies, motivation and learning strategies, methods and activities used in the teaching process. These reasons and solutions can be identified through extensive qualitative research.

REFERENCES

Anderson, P. S. (1987). *Comparison of Student Attitudes About Seven Formats of Educational Testing, With Emphasis on The MDT Multi-Digit Testing Technique.* Paper presented at the Annual Meeting of the Mid-Western Educational Research Association,Chicago.

Bal, P. A. (2012b). Sınıf öğretmenliği öğretmen adaylarının matematik dersine ilişkin ölçme değerlendirme tercihleri. *Türk Eğitim Bilimleri Dergisi*, 10(3), 459-479.

Bal, P. A. (2013). Lisans öğrencilerinin matematik dersine ilişkin değerlendirme tercihleri ile öğrenme stratejileri arasındaki ilişkinin incelenmesi. *International Online Journal of EducationalSciences,5*(1), 242-257.

Bal, P. A. (2012a). Öğrencilerin matematik dersine ilişkin değerlendirme tercihleri. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 27*, 60-73.

Biggs, J. (2003). *Teaching for quality learning at university* (2nd edition). United Kingdom: The Society for Research into Higher Education and Open University Pres.

Biggs, J. B. (1987). *Student approaches to learning and studying.* Melbourne: Australian Council for Educational Research.

Biggs, J. (1999a). What the Student does: Teaching for enhanced learning. *Higher Education*, 18(1), 57-75.

Biggs, J. (1999b). *Teaching for quality learning at university*. London: Open University Press.

Birenbaum, M. (1994). Toward adaptive assessment - the student's angle. *Studies in Educational Evaluation, 20,* 239-255.

Birenbaum, M. (1997). Assessment preferences and their relationship to learning strategies and orientations. *Higher Education*. 33, 71-84.

Birenbaum, M. (2003). New insights into learning and teaching and their implications for assessment. Segers, M., Dochy, F. ve Cascallar, E. (eds.). *Optimizing New Methods of Assessment: In Search of Qualities and Standards* (pp.13-36)*.* Dordrecht, The Netherlands: Kluwer.

Büyüköztürk, Ş. ve Gülbahar, Y. (2010). Assesmentpreferences of highereducationstudents. *EurasionJournal of EducationalResearch*, *41*, 55-72.

Byrne, M., Flood, B. ve Willis, P. (2002). The relationship between learning approaches and learning outcomes: A study of Irish accounting students. *Accounting Education*, *11*(1), 27-42.

Çolak, E. ve Fer, S. (2007). Öğrenme yaklaşımları envanterinin dilsel eşdeğerlik, güvenirlik ve geçerlik çalışması. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 16*(1), 197-212.

Doğan, C. D. (2016). Biçimlendirici değerlendirmenin üniversite öğrencilerinin değerlendirme tercihleri üzerindeki etkisi: Bir ölçekleme çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi,16*(2), 413-431.

Doğan, C. D. (2011). *Öğretmen adaylarının başarıları belirlenirken tercih ettikleri durum belirleme yöntemlerini etkileyen faktörler ve bu yöntemlere ilişkin görüşleri* (Yayınlanmamış Doktora Tezi). Ankara Üniversitesi, Ankara.

Doğan, C. D., Atmaca, S. ve Yolcu, F. A. (2012). Sekizinci sınıf öğrencilerinin öğrenme yaklaşımları ve değerlendirme tercihleri arasındaki ilişki. *İlköğretim Online, 11*(1), 264-272.

Doğan, D. ve Kutlu, Ö. (5-7 Mayıs 2010*). Değerlendirme tercihleri ile öğrenme stratejileri arasındaki ilişki: ingilizce öğretmen adayları örneği*. 2. Ulusal Eğitimde Ölçme ve Değerlendirme Sempozyumunda sunuldu. Mersin Eğitim Fakültesi Mersin. 5-7 Mayıs.Mersin Üniversitesi Eğitim Fakültesi.

Entwistlei, N. J ve Ramsden, P. (1983). Understanding student learning. London: Croom Helm.

Ekinci, N. (2008). *Üniversite öğrencilerinin öğrenme yaklaşımlarının belirlenmesi ve öğretme öğrenme süreci değişkenleri ile ilişkisi*. (Yayımlanmamış Doktora Tezi). Hacettepe Üniversitesi, Ankara.

Gelbal, S. ve Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi,33*, 135-147.

Gijbels, D. ve Dochy, F. (2006). Students' assessment preferences and approaches to learning: can formative assessment make a difference?. *Educational Studies, 32*(4), 399-409.

Guilford, J. P. (1954). *Psychometric methods*. New York: Mc Graw-Hill.

Karasar, N. (2014). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın Dağıtım.

Kember, D., Biggs, J. ve Leung, D. Y. P. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the learning process questionnaire. *British Journal of Educational Psychology*, *74*, 261-280.

Kuzgun, Y. ve Deryakulu, D. (2004). *Bireysel farklılıklar ve eğitime yansımaları.* (Editörler: Y. Kuzgun ve D. Deryakulu). Ankara: NobelYayın-Dağıtım.

Marton, F. ve Saljo, R. (1976a). On qualitative differences in learning I: -Outcome & process. *British Journal of Educational Psychology*, *46*, 4-11.

Marton, F. ve Saljo, R. (1976b). On qualitative differences in learning -II: Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology, 46*, 115-127.

Mayya, S., Rao, A.K. ve Ramnarayan, K. (2004). Learning approaches, learning difficulties and academic performance of undergraduate students of physiotherapy. *The Internet Journal of Allied Health Sciences and Practice, 2*(4), 1-6.

McManus, I.C., Richards, P. ve Winder, B.C. (1999). Intercalated degrees, learning styles, and career preferences: Prospective longitudinal study of UK medical students. *British Medical Journal.International edition, 319*, 542-546.

Phillips, F. (1999). Business students' learning preferences and associated task performance. *Journal of Education for Business,75*(1), 27-32.

Reid, W. A., Duvall, E. ve Evans, P. (2007). Relationship between assessment results and approaches to learning and studying in year two medical students. Medical Education, 41(8), 754-762.

Sambell, K., McDowell, L. ve Brown, S. (1997). 'But is it fair?': An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation, 23*(4), 349-371.

Scouller, K. (2000). The influence of assessment method on student learning. *Australian Association for Research in Education, 35*(4), 453-472.

Struyven, K., Dochy, F. ve Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education,30*(4), 325-341.

Thurstone, L. L. (1927). A law of comparativejudgement. *PsychologicalRewiev*, *34*, 273-286.

Turgut, M. F. (1992). *Eğitimde ölçme ve değerlendirme metotları*. (8. Basım). Ankara: Saydam Matbaacılık.

Zeinder, M. (1987). Essay versus multiple-choice type classroom exams: The student's perpective. *Journal of Education Research,80*, 352-358.