



Student Performance Task Assessment Using Multiple Criteria Decision Making (MCDM) Techniques: An Application for 9th Grade Chemistry Course

Mehmet YÜKSEL*^a, Ömer GEBAN^b

Article Info

DOI: 10.14686/buefad.400787

Article History:

Received 02.03.2018
Accepted 30.07.2018
Published 31.10.2018

Keywords:

Measurement and assessment,
Performance task,
Multiple-Criteria Decision.

Article Type:

Research article

Abstract

This study was conducted with the aim to analyze some of the problems related to measurement and evaluation encountered in student performance task assessment. Multiple criteria decision-making methods were utilized in this study. The criteria adopted in the assessment of performance tasks and the weights of learning domains were calculated through AHP. Predicated upon the weights of cognitive, affective, and psychomotor domains, the criteria were selected through TOPSIS technique. According to the findings of the study, the order of the relative weights of the learning domains considered in the selection of performance criteria was cognitive, affective, and psychomotor respectively. It was also ascertained that the weights of performance criteria groups were subsequently the features of elements, the process of content formation and presentation. The student performance tasks prepared in a 9th grade chemistry course subject titled 'Elements' were assessed in accordance with the techniques and approach suggested by the chemistry teacher in the expert group. It was determined that the score for student performance task was 81.2 (100) within the context of application in this study.

Öğrenci Performans Görevinin Çok Ölçütlü Karar Verme (ÇÖKV) Teknikleri İle Değerlendirilmesi: 9. Sınıf Kimya Dersine Yönelik Bir Uygulama

Makale Bilgisi

DOI: 10.14686/buefad.400787

Makale Geçmişi:

Geliş 02.03.2018
Kabul 30.07.2018
Yayın 31.10.2018

Anahtar Kelimeler:

Ölçme ve değerlendirme,
Performans görevi,
Çok ölçütlü karar verme.

Makale Türü:

Araştırma makalesi

Öz

Bu araştırma öğrenci performans görevinin değerlendirilmesinde karşılaşılan bazı ölçme ve değerlendirme problemlerinin çözümlenmesi amacıyla yapılmıştır. Çalışmada çok ölçütlü karar verme teknikleri kullanılmıştır. Performans görevlerinin değerlendirilmesinde kullanılan kriterler ve öğrenme alanlarının ağırlıkları AHP ile hesaplanmıştır. Kriterlerin seçimi TOPSIS tekniğiyle bilişsel, duyuşsal ve devinişsel alanların ağırlıkları esas alınarak yapılmıştır. Çalışmanın sonuçlarına göre performans kriterlerinin seçiminde dikkate alınan öğrenme alanlarının nispi ağırlıkları sırasıyla bilişsel, duyuşsal ve devinişsel şeklindedir. Performans kriter gruplarının ağırlıkları ise sırasıyla elementin özellikleri, içerik, hazırlık süreci ve sunu yapma olarak belirlenmiştir. 9. sınıf kimya dersinde elementler konusunda öğrenci tarafından hazırlanan performans görevi uzman gurupta yer alan kimya öğretmeni tarafından önerilen teknikler ve yaklaşımla değerlendirilmiştir. Bu çalışmanın uygulaması kapsamında değerlendirilen öğrencinin performans görevi puanının 81.2 (100) olduğu saptanmıştır.

*Corresponding Author: yukselmehmet@gazi.edu.tr

^aDr., Gazi University, Ankara/Turkey, <http://orcid.org/0000-0003-0124-1992>

^bProf. Dr., Middle East Technical University, Ankara/Turkey, <http://orcid.org/0000-0002-9433-0056>

Introduction

It is known that measurement and evaluation are decisive in making sense of the phenomena and events related to different areas of life. Measurement and evaluation also have a significant function in the process of education and teaching activities. Despite the various definitions based on context and fields in the literature, measurement can be defined as the determination or gradation of the qualitative or quantitative properties of phenomena and events. On the other hand, the concept of evaluation is, eventually, the process of decision-making. The process of measurement aims to describe the phenomena and events exactly as they are within the bounds of possibility. However, the results of measurement are compared with the determined criteria in the process of evaluation. In other words, the result of measurement is surveyed whether it corresponds to the condition determined by the criteria. From this aspect, evaluation is different than measurement since it results in a decision on the measured property (Özçelik, 1992). This definition hints a significant relationship between the processes of measurement and evaluation. The fact that evaluation is made only through a completed measurement activity lies in the essence of the relationship between measurement and evaluation. Hence, in its broadest sense, the evaluation of an activity depends primarily on the measurement activity. It can be seen that the evaluation of an activity can be carried out through the measurement process.

In the process of education and teaching, students' academic success is attempted to be determined by various measurement and evaluation approaches (Gelbal & Kelecioğlu, 2007; Güler, 2017; Metin & Özmen, 2010; Yaşar & Sözbilir, 2013; Yıldız & Uyanık, 2004). However, the measurement and evaluation approaches used in the process of education and teaching have several consequences beyond determining students' academic success. Considering within the individual context, it primarily aims to facilitate students' life following the education process through behaviors, skills and knowledge to be acquired. Furthermore, education and teaching activities have a role and significance in the realization of institutional and social objectives. Along with the fact that the realization of micro and macro objectives of education either socially or institutionally is volitional, the level of achieving volitions can be determined by the determination of what level these objectives are attained. This is achieved by measurement and evaluation activities that have a significant functional role in the process of education. Along with the fact that a measurement and evaluation activity takes place in all educational processes involving learning in the progress within the historical process, it is observed that measurement and evaluation approaches differ according to the context fundamental to epistemological theories. To exemplify, it is assumed that the idea that it would not be sufficient to measure students' academic success through evaluation based on merely on their answers in exams was adopted according to the structuralism theory which establishes the fundamental philosophy of Turkish secondary education system. In addition, the idea of evaluating individual and group performances of students was also adopted (Baki & Birgin, 2002; Birgin, 2008). In this respect, there are various measurement and evaluation approaches used in practice (Arı, 2010; Deniz & Kaptan, 2001; Döş, 2016; Kaya, Karaçam, Eş & Tuncel, 2013; Şenel Çoruhlu, Er Nas & Çepni, 2009; Yiğit & Kırımlı, 2015) and examined within theoretical context recently. As a matter of fact, the related literature has numerous methods, approaches and techniques (Altun & Gelbal, 2014; Büyüktokatlı & Bayraktar, 2014; Nazlıççek & Akarsu, 2008) developed for the solution of issues in measurement and evaluation in educational activities. However, one of the fundamental reasons for the difference in measurement and evaluation methods is the variety of topics and academic disciplines subject to education. Another reason can be the fact that a standardized measurement and evaluation approach has not yet been developed for the determination of the effectiveness of education.

In the review of literature, it is seen that measurement and evaluation approaches were examined under two categories (Akbaş & Gençtürk, 2013; Deniz & Kaptan, 2011; Kuran & Kanathı, 2009; Yıldız & Uyanık, 2004). The first one is the traditional measurement and evaluation approach which fundamentally aims to determine the level of student outcomes achieved in the process of education. Traditional measurement and evaluation approaches aim to assess learning on cognitive level. However, the idea that the assessment of learning within affective and psychomotor domains enables the evaluation of meaningful learning is also advocated (Bertiz & Uluçınar Sağır, 2014). As a matter of fact, not only cognitive learning abilities of students but also their development of affective and psychomotor domains were also considered in the learning process according to the view held recently (Demirbaş & Yağbasan, 2004). It can be stated that a differentiation of objectives, in other words, a learning process which aims the development and acquisition of cognitive, affective and psychomotor behaviors might be difficult to evaluate through traditional measurement and evaluation methods. However, research indicates that alternative measurement and evaluation methods are not functional due to a number of limitations and difficulties

including crowded classrooms, time management, the development, implementation and evaluation of measurement instruments, teachers' lack of knowledge and skills regarding alternative measurement and evaluation methods, determination of criteria to assess student tasks, scale development and utilization (Anıl & Acar, 2008; Arı, 2010; Büyüktokatlı & Bayraktar, 2014; Çiftçi, 2010; Dağhan & Akkoyunlu, 2014; Döş, 2016; Gelbal & Kelecioğlu, 2007; Kaya, Karaçam, Eş & Tuncel, 2013; Kuran & Kanatlı, 2009; Metin, 2013; Metin & Özmen, 2010; Nazlıçiçek & Akarsu, 2008; Özdemir, 2010; Sağlam-Arslan, Devocioğlu-Kaymakçı & Arslan, 2009; Şenel Çoruhlu, Er Nas & Çepni, 2009; Yaşar & Sözbilir, 2013).

Research indicates that performance tasks (Altun & Gelbal, 2014; Anıl & Acar, 2008; Arı, 2010; Çiftçi, 2010; Dağhan & Akkoyunlu, 2014; Kaya, Karaçam, Eş & Tuncel, 2013; Metin & Özmen, 2010; Özdemir, 2010; Şenel Çoruhlu, Er Nas & Çepni, 2009; Yiğit & Kırımlı, 2015) as a type of alternative measurement and evaluation approaches are mostly embraced and employed by teachers in Turkish education system. One of the fundamental reasons for this can be the provision in the 50th article of the Regulation on Secondary Education Institutions (MEB, 2013) by the Ministry of National Education (MONE) which covers the regulation on the procedures and principles concerning the operation, administration, teaching and education in formal state and private secondary school institutions affiliated to the MONE in Turkey. As a matter of fact, with the new rearrangement made in 2017 (Article 50- Amendment: Republic of Turkey Official Gazette, 2017) which states that “according to the types of schools, students carry out projects and tasks apart from exams along with seminars, conferences and similar exercises oriented at community service. Students perform at least one performance task in all courses every semester and do one project in at least one course in an academic year.”

It is possible to interpret that the reason for the frequent and common use of performance tasks in practice is through this rearrangement adopted by the related regulation and the MONE. Another reason is the view that performance task assessment is easier to implement than other methods of alternative measurement and evaluation. However, the functionality expected of the approach to the assessment of student performance tasks in the process of education and teaching is possible with the implementation of the conceptual content ascribed to student performance task assessment. This might be possible with the answer to the question of how measurement and evaluation will be made in student performance task assessment. However, there are some problems and limitations to measurement and evaluation encountered in the implementation of performance task assessment approach. As a matter of fact, it can be observed that some problems including the selection of subject to be given as a performance task or project, its appropriateness to students' levels, the number of performance tasks to be given in an academic term, the way assignments are made and assessed, and the selection of assessment criteria which are considered to be serious exist in the implementation of performance task assessment when conducted research in the literature (Arı, 2010; Çiftçi, 2010; Kaya, Karaçam, Eş & Tuncel, 2013; Metin & Özmen, 2010) are examined. When the research on this phenomenon are examined in detail, it can be observed that teachers' lack of knowledge of alternative measurement and evaluation instruments – how the subjects to be assigned as performance tasks will be selected and prepared and how the assessment criteria will be determined and evaluated – the fact that they have problems in time and practice are some of the teacher-induced problems. Individual differences among students, positive or negative attitudes towards performance tasks and projects, much attention attached to the success they will have in the central placement and selection exams lead to problems in the preparation, presentation and execution of their performance tasks. The economic conditions of families, their attention to students, their negative attitudes towards projects and performance tasks, their lack of knowledge of projects and performance tasks of their children, their lack of attention and counsel while their children carry out their projects and performance tasks, too much attention displayed contrarily by some parents, and preparation of performance tasks by parents instead of students are some of parent-induced problems. In general, time, resources, materials, schools' physical capacity, classroom population, lack of exemplary implementations, and practitioners' lack of knowledge are some of the problems encountered in the implementation of alternative measurement and evaluation methods (Anıl & Acar, 2008; Arı, 2010; Çiftçi, 2010; Dağhan & Akkoyunlu, 2014; Kaya, Karaçam, Eş & Tuncel, 2013; Metin & Özmen, 2010; Özdemir, 2010; Şenel Çoruhlu, Er Nas & Çepni, 2009).

The issues discussed above indicate that the evaluation of student success with an objective and reliable approach in performance task assessment is unsatisfactory. A systematic measurement and evaluation approach to student performance task assessment has not been encountered in the review of literature. Minimizing the problems encountered in student performance task assessment by teachers can be possible with the development of the frame of measurement and evaluation of student performance task assessment. The conceptual frame of performance

task assessment requires the implementation of measurement and evaluation to be integrative-quantitative. Furthermore, there are certain detailed and problem-causing issues to be answered in the implementation of performance task assessment. The first of these is that there must be criteria determined in performance task assessment since the measure must be appropriate to the evaluation to reach a correct conclusion in evaluation as stated by Özçelik (1992). Every measure determines the condition for the usefulness of a measurement's result with a different purpose. A conformable measure to be used to make an evaluation with a definite purpose must be sought (Özçelik, 1992). As a matter of fact, it was stated in a study (Metin & Özmen, 2010) that it is important on what criteria to evaluate performance tasks assigned to students by teachers. Another issue encountered in the evaluation of student performance tasks is probably the question of how to evaluate when there is more than one evaluation criterion. Considering this issue is vital to make an objective evaluation of student performance task. In addition to this, there will be numerous criteria in the evaluation of student performance task and a variety of scientific fields and subjects within a performance task. For example, the performance task assessment undertaken within a mathematics course offered in high school education level will differ from the criteria of the performance task assessment made in a chemistry course in terms of quality and quantity. Besides, such a case might occur in the same scientific field along with the differentiation of subject contents.

For instance, the assessment criteria of performance tasks offered on "Atom and Periodic System" may not be similar in terms of quantity and quality to those of prepared on a subject titled "Chemistry Field Interests" in a chemistry course. Another aspect of the criteria that stems from being a feature of the evaluation of integrative systems (Chen, Hsieh & Do, 2015; Lai & Choi, 2015; Lupo, 2013; Wu, Chen, Chen & Zhuo, 2012) is the relative importance attached to them. In other words, the level of relative importance of the criteria that must be considered in performance task assessment may not be the same as the one in the process of measurement and evaluation. For instance, the relative importance of "n1" which is one of the "n" criteria based in student performance task assessment may be different from those of "n2" or "nn". Therefore, student performance task assessment needs to be carried out in the basis of weight considering the potential differences of importance of the criteria. Another issue encountered in practice is how to measure student performance task according to the determined criteria. For instance, the question for teachers is, based on the criteria, how to and with which instrument to measure the performance task about "The Physical States of Matter" which was a study subject in a performance task prepared in a chemistry course.

Within the context of the views expressed in the introduction part of this study, the second part includes the techniques used in the study along with the theoretical framework of the approach proposed in accordance with the evaluation of student performance tasks. In the third part of the study, the findings of the evaluation of a student performance task on the subject of "Elements" determined as a consequence of course outcomes (MEB, 2017) in the unit of "Science of Chemistry" in accordance with a 9th grade chemistry course were presented. The fourth part includes the evaluation of the results of the study and further suggestions.

Method

The study consists mainly of two stages. In the first stage, the criteria for student performance task assessment were determined and selected while the measurement and evaluation of student performance task assessment based on the determined criteria were presented in the second stage.

In this study, the implementation of the method mentioned above encompasses student performance task assessment of a 9th grade chemistry course subject titled "Elements" in the academic year of 2017-2018. Data were collected in two ways in accordance with the purpose of the study. First, the data of this study, as in the relevant literature of multiple criteria decision-making (Dağdeviren, Yavuz & Kılınc, 2009; Shyur, 2006; Yüksel & Geban, 2015), were gathered from expert opinion in compliance with the methods used in this study. With this purpose in mind, the expert group of this study consists of the first researcher, an expert chemistry teacher working in a high school in Ankara, and the people who bear the title as chemistry teachers. Other members of the expert group were informed by the first researcher about the purpose of the study, student performance task, the criteria for pairwise comparison and the evaluation of the criteria according to learning domains. The expert group made the following pairwise comparisons and evaluations based on common view:

- ✓ Pairwise comparisons of learning domains

- ✓ Evaluation of performance task assessment criteria according to learning domains
- ✓ Pairwise comparisons of main groups of the criteria fundamental to performance task
- ✓ Pairwise comparison of performance task assessment criteria

In the first stage of the implementation, evaluations were made based on expert opinion. Using the scale developed by Saaty (1980) in Table 1, the expert group formed in this study evaluated pairwise comparisons of performance task assessment criteria and pairwise comparisons of main groups of the criteria fundamental to performance task as well as pairwise comparisons according to learning domains.

Table 1. Level of Importance in Pairwise Comparisons

a_{ij}	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Weak importance	An activity is favored very slightly over another
5	Strong importance	An activity is favored strongly over another
7	Very strong or demonstrated importance	An activity is favored very strongly over another
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2,4,6,8	Intermediate values	Used when necessary

In the second stage of the study, the assessment of student performance task was made by the 9th grade chemistry teacher in the expert group. In relation to the evaluations of the findings in the first and second stages, data were analyzed through Expert Choice (2000) and Microsoft Excel (2016) programs.

Multiple Criteria Decision-Making Techniques Used in The Study

TOPSIS (Technique for Order Performance by Similarity to Ideal Solution) and AHP (Analytic Hierarchy Process) which are two of multiple criteria decision-making techniques were utilized in this study about student performance task assessment.

AHP Technique

Developed by Saaty (1980), AHP is a technique utilized frequently in problems related to multiple criteria decision-making in the literature (Gupta & Chaudary, 2017; Lupo, 2013; Wu, Chen & Zhuo, 2012; Zyoude & Fuchs-Hanusch, 2017). The main characteristic of the AHP technique is that the problem of analysis is a whole of the elements consisting of items. Because of this feature, the items that constitute the problem of research primarily in the AHP technique are organized hierarchically according to the objective function. This kind of distinction ensures that all the elements that make up the problem of research are in the model. Thus, pairwise comparison matrices are constructed according to the hierarchical levels and grouping of the items in the generated model (Saaty, 1980; Saaty, 1986). According to this, in the AHP technique, the comparison of i. component and j. component is indicated by a_{ij} . In the parallel model, the pairwise comparison of j. component and i. component is indicated by a_{ji} . In other words, the value of a_{ji} component in AHP is equal to the value of $a_{ji} = 1/a_{ij}$ equation. In pairwise comparisons of the components in AHP model, the scale (Table 1) proposed by Saaty (1980) is most utilized. Another characteristic of AHP technique is the square matrices of nxn size where pairwise comparison matrices are made (n-1)/2 times.

As a result of the pairwise comparisons made in AHP, nxn size square is indicated in the matrix as follows (Saaty, 1980):

$$A = \begin{bmatrix} 1 & a_{12} & \dots & a_{1n} \\ 1/a_{12} & 1 & \dots & a_{2n} \\ \dots & \dots & 1 & \dots \\ 1/a_{1n} & 1/a_{2n} & \dots & 1 \end{bmatrix}$$

The priority vector is determined by the equation of $(A - \lambda_{max} I) W = 0$. Within this equation, A symbolizes the pairwise comparison matrix, W eigenvector and λ_{max} the highest eigenvalue of A matrix. Another characteristic of AHP is the ability to determine how inconsistent the pairwise comparisons concerning the evaluation of superiority are (Saaty, 1980; Saaty, 1991). It is necessary to calculate the consistency index (C.I.) and random index (R.I) to determine the inconsistency ratio (I.R.) The consistency index is calculated by the equation of $C.I = (\lambda_{max} - n) / (n - 1)$. In the equation, n symbolizes the dimension of the matrix. The value of random index, on the other hand, is determined by the dimension of the matrix (Table 2). Accordingly, the inconsistency ratio of pairwise comparison matrix is calculated by the equation of $C.R = C.I / R.I$. Even though the matrix can be different based on its dimension, the consistency of pairwise comparisons is possible, with C.R. being less than 0.10. Pairwise comparisons are remade when the inconsistency ratio is larger than 0.10 in AHP. Pairwise comparisons of the components in the scale and the calculation of the priorities and consistency ratios are performed for grouping at all hierarchical levels in the scale.

Table 2. Random Index Values (RI) (Saaty, 1980)

(n)	1	2	3	4	5	6	7	8	9
(RI)	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45

TOPSIS Technique

TOPSIS, which is another method for multiple criteria decision-making, enables the evaluation and analysis of numerous alternatives within the context of more than one criterion. TOPSIS was preferred in this study since it was believed that it would be the most appropriate approach to determine the criteria to be used in student performance task assessment and that one of the principal characteristics of TOPSIS technique is that it enables the order of choices within the context of more than one criterion. In this study, TOPSIS technique was used since it was believed that the criteria selection to be used in student performance task assessment might be under more than one dimension. TOPSIS, one of the multiple criteria decision-making techniques, was first developed by Hwang and Yoon (1981). It was used in the studies aimed to resolve multiple criteria problems through TOPSIS technique (Dağdeviren, Yavuz & Kılınç, 2009; Ersöz, Kabak & Yılmaz, 2011; Zyoud & Fuchs-Hanusch, 2017). The main steps of TOPSIS technique are as follows (Shyur, 2006):

Step 1: *Form the decision matrix for order.* The structure of the decision matrix is illustrated as below. A_i shows the potential alternatives; $i = 1, \dots, m$; F_i indicates the related i . characteristic or criterion, and $j = 1, \dots, n$; f_{ij} is a value on condition that each alternative is compared to A_i and each F_j criterion.

$$D = \begin{matrix} & \begin{matrix} F_1 & F_2 & \dots & F_j & \dots & F_n \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_j \end{matrix} & \begin{bmatrix} f_{11} & f_{12} & \dots & f_{1j} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2j} & \dots & f_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ f_{i1} & f_{i2} & \dots & f_{ij} & \dots & f_{in} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ f_{j1} & f_{j2} & \dots & f_{jj} & \dots & f_{jn} \end{bmatrix} \end{matrix}$$

Step 2: *Calculate the standardized decision matrix.* $R (= [r_{ij}])$. The normalized value r_{ij} is calculated by the following equation:

$$r_{ij} = \frac{f_{ij}}{\sqrt{\sum_{j=1}^n f_{ij}^2}}, j = 1, \dots, n; i = 1, \dots, m.$$

Step 3: *Calculate the weighted normalized decision matrix* which is calculated by the multiplication of normalized decision matrix and the related weights. Weighted normalized values v_{ij} are calculated as follows: the $v_{ij} = w_j r_{ij}$, $j = 1, \dots, n; i = 1, \dots, m$, w_j statement indicates the weight of j . characteristic or criterion.

Step 4. Identify the ideal and negative ideal solutions. J indicates benefit criterion and J' cost criterion, it follows as:

$$V^+ = \{v_1^+, \dots, v_n^+\} = \{(\max_i v_{ij} | j \in J), (\min_i v_{ij} | j \in J')\}$$

$$V^- = \{v_1^-, \dots, v_n^-\} = \{(\min_i v_{ij} | j \in J), (\max_i v_{ij} | j \in J')\},$$

Step 5. Calculate split measurements. This is calculated using m-dimensional Euclidean distance. Splitting each alternative from the ideal solution (D_i^+) is as follows:

$$D_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2}, i = 1, \dots, m, \text{ likewise splitting each alternative from the negative ideal solution}$$

(D_i^-) is as follows:

$$D_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}, i = 1, \dots, m$$

Step 6: Calculate the relative distance to the ideal solution and identify the order of choices.

The following equation demonstrates how the calculation and ordering are done: $C_i = \frac{D_i^-}{D_i^+ + D_i^-}, i = 1, \dots, m.$

C_i index value gets a value between 0 and 1. Large index values show better performance alternatives.

Findings

In this part of the study, the findings of student performance task assessment within the context of 'Elements' in a 9th grade chemistry course was introduced.

In the first step of this study, the criteria formation and selection based in student performance task assessment was done. For this, firstly the criteria of student performance task assessment were determined. In the related literature, no criterion was encountered in relation to the evaluation of a chemistry course or content within the context of the specific chemistry subject. However, it was observed that some general performance task assessment criteria were available in Secondary Education Chemistry Course Syllabus (MEB, 2007) and Secondary Education 9th Grade Chemistry Course Book (Dursun, Gülbay, Çetin & Tek, 2012; Güntut, Güneş & Çetin, 2017) and suggested to be used in Turkish education system. In addition to the fact that these resources (Dursun, Gülbay, Çetin & Tek, 2012; Güntut, Güneş & Çetin, 2017; MEB, 2007) were taken as a basis in the determination of the criteria fundamental to this study, the criteria representing the course subject titled 'Elements' were developed in this study (Table 3).

Table 3. Student Performance Task Assessment Criteria

Code	Criteria
P01	Ability to present descriptive information
P02	Ability to present physical properties
P03	Ability to present chemical properties
P04	Ability to demonstrate its position in periodic system
P05	Ability to present exploratory information
P06	Ability to explain its uses in daily life
P07	Ability to introduce basic use of area in industry

P08	Ability to give explanations of its effects on human health
P09	Ability to give explanations of its effects on the environment
P10	Ability to set the objective for performance tasks
P11	Ability to make a study plan applicable to performance
P12	Ability to identify the necessary resources
P13	Ability to use multiple and various subject-related resources
P14	Ability to direct relation of resources to the subject
P15	Up-to-date information
P16	Ability to use subject-related principles, phenomena and concepts appropriately
P17	Ability to support comments with proofs in resources
P18	Ability to use Turkish grammar and spelling rules
P19	Ability to display of critical thinking ability
P20	Ability to use creativity skills
P21	Ability to organize collected data through analysis
P22	Ability to implement performance tasks according to plan
P23	Ability to speak Turkish fluently and accurately
P24	Ability to support presentations with target-specific materials
P25	Ability to present topic by attracting the audience's attention
P26	Ability to respond to questions
P27	Ability to use fluent and body language
P28	Ability to manage time effectively during presentation
P29	Ability to possess self-confidence during presentation

The criteria in Table 3 were subject to a two-step analysis. In the first step, 29 criteria about the course subject titled 'Elements' included in the 9th grade chemistry course content and considered to be sufficient were determined. In the second step, TOPSIS method was utilized to determine the degree of importance to which those 29 criteria measure a performance task designed on a course subject titled 'Elements'. The weights of the dimensions to be taken as a basis in the criteria selection of student performance task assessment were calculated by AHP method. Cognitive, affective and psychomotor learning domains were taken as a basis in the determination of the degree of importance in the evaluation of the criteria of performance task in this study. Accordingly, the formed pairwise comparison matrix and the calculated local weights were presented in Table 4. The consistency ratio (CR= 0.05) of the pairwise comparison matrix was found to be satisfactory in the calculation made.

Table 4. The Weights and Pairwise Comparisons of Learning Domains

Dimensions of Learning Domains	CD	AD	PD	Weights
Cognitive (CD)	1	2	2	0.493
Affective (AD)		1	2	0.311
Psychomotor (PD)			1	0.196

The criteria for the assessment of the performance task in the study were weighted according to the learning areas (cognitive, emotional and psychomotor dimensions), and then the number of criteria was reduced by TOPSIS technique according to importance level. This is due to the fact that fewer but valid criteria are used to assess the performance task, so that the performance task evaluation process is objective and facilitated. As a result of the TOPSIS analysis (Table 5), the criteria to be based on performance evaluation were determined considering the importance level. The arithmetic mean and the standard deviation value are taken into account as the critical values for determining the criteria based on performance task evaluation. Accordingly, the criteria one standard deviation ($-\sigma$) under the arithmetic average (μ) of the order values calculated through TOPSIS analysis were chosen. Thus, the criteria within the interval of 84.1% [$0.50+ (1\sigma)$] were chosen as far as the area under the normal distribution curve was concerned. To do this, initially the arithmetic average (3.0267586) and standard deviation (0.4320363) of the degrees of importance were calculated through TOPSIS analysis. The critical value was found to be ($\mu-\sigma$) 2.594722. The criteria that have larger degrees of importance than this critical value were taken as a basis (Table 5).

Table 5. Evaluation of the Criteria of TOPSIS Analysis according to Learning Domains

Code	Criteria	Cogni- tive 0.493	Affec- tive 0.311	Psycho- motor 0.196	Order value
P01	Ability to present descriptive information	5	3	2	3.790
P02	Ability to present physical properties	5	1	2	3.168
P03	Ability to present chemical properties	5	1	2	3.168
P04	Ability to demonstrate its position in periodic system	5	3	2	3.790
P05	Ability to present exploratory information	4	3	3	3.493
P06	Ability to explain its uses in daily life	4	3	3	3.493
P07	Ability to introduce basic use of area in industry	4	3	2	3.297
P08	Ability to give explanations of its effects on human health	4	3	3	3.493
P09	Ability to give explanations of its effects on the environment	4	3	3	3.493
P10	Ability to set the objective for performance tasks	4	2	2	2.986
P11	Ability to make a study plan applicable to performance	4	2	2	2.986
P12	Ability to identify the necessary resources	4	2	3	3.182
P13	Ability to use multiple and various subject-related resources	4	2	2	2.986
P14	Ability to direct relation of resources to the subject	3	2	2	2.493
P15	Up-to-date information	3	2	2	2.493
P16	Ability to use subject-related principles, phenomena and concepts appropriately	4	3	2	3.297
P17	Ability to support comments with proofs in resources	3	3	2	2.804
P18	Ability to use Turkish grammar and spelling rules	3	3	3	3.000
P19	Ability to display of critical thinking ability	3	3	3	3.000
P20	Ability to use creativity skills	3	3	3	3.000
P21	Ability to organize collected data through analysis	3	2	3	2.689
P22	Ability to implement performance tasks according to plan	3	3	3	3.000
P23	Ability to speak Turkish fluently and accurately	3	2	3	2.689
P24	Ability to support presentations with target-specific materials	3	3	3	3.000
P25	Ability to present topic by attracting the audience's attention	3	2	2	2.493
P26	Ability to respond to questions	3	3	3	3.000
P27	Ability to use fluent and body language	2	1	2	1.689
P28	Ability to manage time effectively during presentation	3	3	3	3.000
P29	Ability to possess self-confidence during presentation	3	3	2	2.804

As can be seen in Table 5, the 25 criteria determined according to degree of importance following TOPSIS analysis were classified under four categories (Table 6) and these criteria were recoded. According to the classification in Table 6, pairwise comparison of primary and secondary criteria was made through AHP technique (Table 6, Table 7). Following the pairwise comparisons, initially the weights of the criteria and then those of main groups were calculated through Expert Choice (2000) program. It was found that the consistency ratio of pairwise comparison matrix (Table 6) of the main groups was (CR) 0.09 and this ratio indicated consistent pairwise comparisons.

Table 6. Pairwise Comparison of the Main Groups of the Criteria Fundamental to Performance Tasks

Groups	E	P	C	P
The properties of elements (E)	1	5	3	7
Preparation process (P)		1	1/3	5
Content (C)			1	5
Presentation(P)				1

Table 7. Pairwise Comparison of Performance Task Assessment Criteria

Criteria											
THE PROPERTIES OF ELEMENTS (E)		E1	E2	E3	E4	E5	E6	E7	E8	E9	
E1	Ability to present descriptive information (E1)	1	3	3	3	5	5	5	5	5	
E2	Ability to present physical properties (E2)		1	1	1/3	3	3	3	3	3	
E3	Ability to present chemical properties (E3)			1	1/3	3	3	3	3	3	
E4	Ability to demonstrate its position in periodic system (E4)				1	3	3	3	3	3	
E5	Ability to present exploratory information (E5)					1	1/3	1/3	1/3	1/3	
E6	Ability to explain its uses in daily life (E6)						1	3	1	1	
E7	Ability to introduce basic use of area in industry (E7)							1	1/3	1/3	
E8	Ability to give explanations of its effects on human health (E8)								1	3	
E9	Ability to give explanations of its effects on the environment (E9)									1	
PREPARATION PROCESS (P)		P1	P2	P3	P4						
P1	Ability to set the objective for performance tasks (P1)	1	3	3	3						
P2	Ability to make a study plan applicable to performance (P2)		1	1/3	1/3						
P3	Ability to identify the necessary resources (P3)			1	1						
P4	Ability to use multiple and various subject-related resources (P4)				1						
CONTENT (C)		C1	C2	C3	C4	C5	C6	C7			
C1	Ability to use subject-related principles, phenomena and concepts appropriately (C1)	1	1	3	3	3	1	3			
C2	Ability to support comments with proofs in resources (C2)		1	3	3	3	1	3			
C3	Ability to use Turkish grammar and spelling rules (C3)			1	1/3	1/3	1/3	1/3			
C4	Ability to display of critical thinking ability (C4)				1	1	1	3			
C5	Ability to use creativity skills (C5)					1	1	3			
C6	Ability to organize collected data through analysis (C6)						1	3			
C7	Ability to implement performance tasks according to plan (C7)								1		
PRESENTATION (P)		P1	P2	P3	P4	P5					
P1	Ability to speak Turkish fluently and accurately (P1)	1	1/3	1/3	3	3					
P2	Ability to support presentations with target-specific materials (P2)		1	1	3	3					
P3	Ability to respond to questions (P3)			1	3	3					
P4	Ability to manage time effectively during presentation (P4)				1	1					
P5	Ability to possess self-confidence during presentation (P5)					1					

In the second step of the study, student performance task assessment was performed based on the criteria and calculated weights. Student performance tasks were evaluated according to the criteria presented in Table 8. Consisting of six grades, this scale was used in multiple criteria decision-making problems in the literature (Yüksel & Dağdeviren, 2006; Yüksel, 2011). The values of the evaluation scale ranged between 0.0 and 1.0.

Table 8. Evaluation Scale

Levels of Evaluation	Level value
Very good (VG)	1.0
Good (G)	0.8
Neutral (N)	0.6
Unsatisfactory (U)	0.4
Very unsatisfactory (VU)	0.2
Not assessable (NA)	0.0

In this study, the assessment of student performance task prepared in a course subject titled ‘Elements’ was performed according to the determined criteria in Table 8. Based on each criterion presented in the second column of Table 9, the evaluation of student performance task was made according to the levels of Table 8. The level of evaluation was presented in the fifth column of Table 9 while the value corresponding with the level was provided in the sixth column. Students’ performance levels based on each criterion were presented in the last column of Table 9. This value is comprised of the multiplication of the value in the sixth column and the general weight value of each criterion. In the last line of Table 9, the student scores based on the evaluation of all criteria were provided. Within the scope of the implementation of this study, the score of student performance task was determined to be 81.2 (100).

Table 9. The Criteria Weights and Results of Student Performance Task Assessment

Groups and Weights	Criteria	Local Weights	General Weights (gw)	Present State	Value (v)	Performance Level (gw×dv)
The Properties of Elements (E) 0.556	E1	0.299	0.166	VG	1.0	0.166
	E2	0.123	0.069	VG	1.0	0.069
	E3	0.123	0.069	VG	1.0	0.069
	E4	0.186	0.103	VG	1.0	0.103
	E5	0.032	0.018	G	0.8	0.014
	E6	0.063	0.035	N	0.6	0.021
	E7	0.041	0.023	G	0.8	0.018
	E8	0.075	0.041	N	0.6	0.025
	E9	0.058	0.032	N	0.6	0.019
Preparation Process (P) 0.136	P1	0.487	0.066	G	0.8	0.053
	P2	0.096	0.013	G	0.8	0.010
	P3	0.208	0.028	G	0.8	0.022
	P4	0.208	0.028	N	0.6	0.017
Content (C) 0.259	C1	0.233	0.060	G	0.8	0.048
	C2	0.233	0.060	N	0.6	0.036
	C3	0.049	0.013	G	0.8	0.010
	C4	0.124	0.032	N	0.6	0.019
	C5	0.124	0.032	N	0.6	0.019
	C6	0.166	0.043	N	0.6	0.026
	C7	0.069	0.018	G	0.8	0.014
Presentation(P) 0.049	P1	0.175	0.009	G	0.8	0.007
	P2	0.326	0.016	N	0.6	0.009

	P3	0.326	0.016	G	0.8	0.013
	P4	0.086	0.004	G	0.8	0.003
	P5	0.086	0.004	N	0.6	0.002
Student Performance Task Level						0.812

Discussion and Conclusion

Using multiple criteria decision-making approaches, this study aimed to resolve some problems encountered in the assessment of student performance tasks which are one of the alternative methods of measurement and evaluation based on structuralism utilized in secondary education of Turkish education system. When the results of the study were examined theoretically, multiple criteria decision-making approaches enabling the integrative evaluation were utilized since this study aimed to resolve the target problems along with the fact that the concept of performance mostly consists of more than one component. As a matter of fact, in the literature as well, it can be observed that multiple criteria methods were employed in the studies to resolve problems related to criteria selection or that there are more than one criterion (Chen, Hsieh & Do, 2015; Gupta & Chaudary, 2017) or the concept of performance as a research topic (Lai & Choi, 2015). With respect to the discussion made above, the reason for the preference to the methods used in this study is that AHP and TOPSIS have analytical implementation that can be evaluated along with integrative and numerous elements. In fact, the purposefulness of the methods used in this study was in harmony with the results of those in the literature (Chen, Hsieh, Do, 2015; Gupta & Chaudary, 2017; Lupo, 2013; Wu, Chen, Chen & Zhuo, 2012; Zyoued & Fuchs-Hanusch, 2017).

As there is no study, to the best knowledge of the present researchers, exploring student performance task assessment with a systematic and multiple criteria approach in the literature, the results of this study were discussed in itself. It is considered significant for the validity of measurement and evaluation to determine the criteria to be taken as a basis in student performance task assessment with a systematic approach. For this reason, the first of the results of this study which achieved the implementation using the suggested method is the determination of the criteria for performance task assessment. The determination of the criteria for evaluation indicates in which context measurement and evaluation will be made. Hence, a valid and reliable platform will be ensured for what is to be measured or evaluated. Otherwise, it is not expected of student performance task assessment to be meaningful without the determined criteria.

Another result of the study is the reduction of the number of criteria determined to be used in student performance task assessment through TOPSIS technique. The consequence of this operation was the opportunity to implement for teachers to assess student performance tasks with fewer criteria with no time constraints and with ease. As a matter of fact, previous research (Anil & Acar, 2008; Dağhan & Akkoyunlu, 2014; Gelbal & Kelecioğlu, 2007; Kaya et al., 2013) indicated that the difficulty of implementation process and time-taking characteristic were considered to be among the reasons for the disuse of alternative measurement and evaluation methods.

Another important result of this study to the related literature is that the criteria could be determined using neither conclusive nor intuitive approach but objectively through TOPSIS which is one of the multiple criteria decision-making methods. Such a selection of criteria enables the determination of highly-representative criteria in measurement and evaluation through an examination of numerous factors.

Another point to be stressed is that measurement and evaluation must be an approach considering the subject or the areas of curriculum outcomes with an integrative fashion. In this respect, the criteria selection was made on the basis of student outcomes including cognitive, affective and psychomotor aim in this study. Thus, the criteria that had little effect on or no contribution to learning domains were used in measurement and evaluation, which is a consequence that reinforces the validity of measurement and evaluation.

Another characteristic of this study is that the criteria were determined through AHP considering their relative weights. The determination of the relative weights of the criteria has a significant function in terms of the validity of student performance task assessment since it is hard to tell or expect of all the criteria taken as a basis in any student performance task assessment to be the same in terms of level of significance or weight. By their nature, criteria might be different in terms of weight and significance depending on their content. For this reason, on what level the criteria to be used in student performance task assessment were weighted was determined in this study. Additionally, consistency ratios were found and calculations were made to see how reliable the expert opinion was in the determination of criteria weights. In other words, it was ensured that the possible inconsistencies of expert

opinions were not included in the evaluation process. Therefore, student performance task assessment was achieved with a reliable approach.

According to the results of the study, a comprehensive evaluation of student performance task in a course subject titled 'Elements' in 9th grade chemistry course was made with an ultimate aim and student performance task assessment was achieved using 25 sub-criteria collected under four main groups including the properties of elements, preparation process, content and presentation. Student performance task assessment was evaluated holistically and thus the student success (81.2%) was determined. Furthermore, how student performance task performed in relation to each main group and criterion was able to be determined. The fact that performance tasks were evaluated comprehensively not only enabled the determination of incomplete and insufficient aspects of students but also revealed what aspects to develop. Henceforth, the purpose of performance assessment is not only reaching a conclusion but also an opportunity for education objectives to be gained by students as a process.

The results of the study indicated that student performance tasks require measurement and evaluation to be done with a multi-dimensional approach and that this characteristic must be represented in the implementation process. It is also considered that this study, in terms of method and implementation, will contribute to the practice of student performance task in Turkish education system since no systematic framework of evaluation exists for teachers despite the fact that performance tasks are included as an obligatory means of assessment in the present secondary education curriculum. It can be stated that this study may contribute to the development of systematics of performance task assessment through its theoretical aspect and implementation results. For instance, the criteria for student performance assessment for teachers in chemistry courses can be determined using the suggested approach in this study. Considering the determined criteria, teachers may contribute to the objective determination of student performance tasks in chemistry courses. In addition, teachers will be provided with copies of resources from various stakeholders in performance task assessment.

One of the studies that can be conducted following the present study is the measurement and evaluation through fuzzy numbers (Şen, 2001). In this study, definite numbers were used in the selection of performance evaluation criteria, determination of weights and assessment of student performance task on the basis of criteria. As is known, definite numbers are real numbers defined according to the classical cluster concept based on syllogistic logic (Şen, 2001). However, it may not always be possible to define the conditions and situations in the decision making and evaluation problems of definite numbers. For this reason, it is suggested that future studies should be done with fuzzy numbers, since it would be possible to make a more appropriate assessment if the measurements and evaluations are made with a fuzzy approach. For the first time, fuzzy numbers expressed according to the concept of fuzzy clusters defined by the Azerbaijani scientist Lütüf Askerzade on the basis of fuzzy logic theory are expressed as the most appropriate numbers to measure uncertainty (Şen, 2001). Another study that can be done in the future is to improve the validity and reliability dimension of the approach proposed in this study with the studies covering different fields of science such as Mathematics, Physics, Biology, and Geography.

Öğrenci Performans Görevinin Çok Ölçütlü Karar Verme (ÇÖKV) Teknikleri İle Değerlendirilmesi: 9. Sınıf Kimya Dersine Yönelik Bir Uygulama

Giriş

Yaşamın farklı alanlarına ilişkin olgu ve olayların anlamlandırılmasında ölçme ve değerlendirmenin belirleyici olduğu bilinmektedir. Eğitim ve öğretim faaliyeti sürecinde de ölçme ve değerlendirmenin önemli bir işlevi bulunmaktadır. Literatürde bağlam ve alanlarına göre farklı tanımlamaları bulunmakla birlikte en genel anlamda ölçme olgu ve olayların sahip oldukları nitel ya da nicel özelliklerinin belirlenmesi ya da derecelendirilmesi işlemidir. Öte yandan değerlendirme kavramı ise neticede bir karar verme işlemidir. Ölçme işleminde olgu ve olayın olanaklar çerçevesinde aslına uygun olarak betimlenmesine çalışılmaktadır. Oysaki değerlendirme işleminde ise ölçme sonuçları belirlenmiş olan ölçütlerle karşılaştırılmaktadır. Diğer bir ifadeyle ölçme sonucu ölçütle belirlenen koşulu karşılayıp karşılamadığına bakılmaktadır. Değerlendirme bu yönüyle ölçme işleminden farklılık göstermekte ve değerlendirme, ölçülen özelliğe ilişkin bir kararla sonuçlanmaktadır (Özçelik, 1992). Bu tanımlama ölçme ve değerlendirme işlemleri arasında önemli bir ilişkinin bulunduğunu ifade etmektedir. Ölçme ve değerlendirme kavramları arasındaki ilişkinin özünde değerlendirmenin yapılabilmesinin ancak ölçme faaliyetinin yapılmış olması bulunmaktadır. Dolayısıyla en genel anlamda bir etkinliğin değerlendirilmesi öncelikle yapılacak olan ölçme işlemine bağlıdır. Ölçme işleminin gerçekleştirilmesiyle faaliyetin değerlendirilmesinin yapılabildiği görülmektedir.

Eğitim ve öğretim sürecinde öğrencinin akademik başarısı çeşitli ölçme ve değerlendirme yaklaşımlarıyla belirlenmeye çalışılmaktadır (Gelbal ve Kelecioğlu, 2007; Güler, 2017; Metin ve Özmen, 2010; Yaşar ve Sözbilir, 2013; Yıldız ve Uyanık, 2004). Ancak eğitim ve öğretim sürecinde kullanılan ölçme ve değerlendirme yaklaşımının öğrencinin akademik başarısını belirlemenin ötesinde çeşitli sonuçları bulunmaktadır. Bireysel bağlamda düşünüldüğünde öncelikle öğrencinin edinmiş olacağı bilgi, beceri ve davranışları ile eğitim süreci sonrasındaki yaşamını kolaylaştırmayı amaçlamaktadır. Bununla birlikte eğitim ve öğretim faaliyeti kurumsal ya da toplumsal amaçların gerçekleşmesine yönelik bir faaliyet olarak önemi ve rolü bulunmaktadır. Toplumsal ya da kurumsal açıdan eğitimin mikro ya da makro amaçlarının gerçekleşmesi öncelikli bir istem olmakla birlikte, bu amaçların ne düzeyde gerçekleşmiş olduğunun belirlenmeye çalışılmasıyla istemlere ulaşma düzeyi anlaşılabilir. Bu da eğitim sürecinde önemli bir işlevsel rolü bulunan ölçme ve değerlendirme faaliyetiyle yapılmaktadır. Tarihsel süreç içerisindeki gelişim öğrenme faaliyetinin olduğu her eğitim sürecinde bir ölçme ve değerlendirme işlemi olmakla birlikte, epistemolojik kuramların temel aldığı bağlama göre ölçme ve değerlendirme yaklaşımlarının farklılaştığı görülmektedir. Örneğin ortaöğretim düzeyindeki Türk eğitim sisteminin temel felsefesini oluşturan yapısalcı bilgi kuramına göre, yalnızca öğrencinin sınavlardaki cevaplarına göre bir değerlendirmenin öğrencinin akademik başarısını ölçmede yeterli olamayacağı görüşünün benimsendiği varsayılmaktadır. Bununla birlikte öğrencinin bireysel ve grup olarak gösterdiği performansın da değerlendirilmesi görüşü benimsenmiştir (Baki ve Birgin, 2002; Birgin, 2008). Bu bağlamda son dönemde kuramsal bağlamda incelenen ve uygulamada kullanılan çeşitli ölçme ve değerlendirme yaklaşımları bulunmaktadır (Arı, 2010; Deniz ve Kaptan, 2011; Döş, 2016; Kaya, Karaçam, Eş ve Tuncel, 2013; Şenel Çoruhlu, Er Nas ve Çepni, 2009; Yiğit ve Kırımlı, 2015;). Nitekim ilgili literatürde eğitim faaliyetinde ölçme ve değerlendirme meselesinin çözümüne yönelik geliştirilen çok sayıda yöntem, yaklaşım ve teknik bulunmaktadır (Altun ve Gelbal, 2014; Büyüktokatlı ve Bayraktar, 2014; Nazlıçipek ve Akarsu, 2008). Bununla birlikte ölçme ve değerlendirme yöntemlerindeki farklılığın başlıca nedenlerinden bir diğeri eğitime konu olan akademik alanların ve konuların çeşitliliğidir. Bir diğer neden ise eğitimin etkililiğini belirlemeye yönelik standart bir ölçme ve değerlendirme yaklaşımının henüz geliştirilememiş olması söylenebilir.

Literatür incelemesinde ölçme ve değerlendirme yaklaşımlarının iki kategoride incelendiği görülmektedir (Akbaş ve Gençtürk, 2013; Deniz ve Kaptan, 2011; Kuran ve Kanatlı, 2009; Yıldız ve Uyanık, 2004;). Bunlardan ilki olan geleneksel ölçme ve değerlendirme yaklaşımında temel amaç öğrencinin eğitim sürecinde edinmiş olduğu kazanımların ne düzeyde gerçekleşmiş olduğunun belirlenmesine yöneliktir. Geleneksel ölçme ve değerlendirme yaklaşımları, öğrenmeyi bilişsel düzeyde ölçmeye çalışmaktadır. Oysa öğrenmede duyuşsal ve psikometri alanlarındaki öğrenmelerin de ölçülmesinin anlamlı olacağı görüşü de ileri sürülmektedir (Bertiz ve Uluçınar Sağır, 2014). Nitekim son dönemde kabul edilen görüş olarak öğrencilerin yalnızca bilişsel öğrenmelerinin yeterliliği değil bununla birlikte duyuşsal ve psikomotor alanlarının da gelişiminin öğrenme sürecinde dikkate alındığı görülmektedir (Demirbaş ve Yağbasan, 2004). Eğitimde böyle bir amaç farklılaşmasının diğer bir ifadeyle bilişsel, duyuşsal ve psikomotor davranışlarının kazandırılması ve geliştirilmesini amaçlayan bir öğrenme sürecinin geleneksel ölçme ve değerlendirme yöntemleriyle değerlendirilmesinin güç olduğu söylenebilir. Ancak alternatif ölçme ve değerlendirme yöntemlerinin uygulamasında sınıfların kalabalık olması, zaman problemi, ölçme araçlarının hazırlanması, uygulanması, değerlendirilmesi, öğretmenlerin alternatif ölçme yöntem ve teknikleri hakkında yeterli bilgi

ve beceriye sahip olmamaları, öğrenci çalışmalarını değerlendirmek amacıyla kriter belirleme, ölçek hazırlama, ölçekleri kullanma gibi bir dizi kısıtlılık ve güçlüklerden dolayı işlevsel olmadığı yapılan araştırmalarda görülmektedir (Anıl ve Acar, 2008; Arı, 2010; Büyüktokatlı ve Bayraktar, 2014; Çiftçi, 2010; Dağhan ve Akkoyunlu, 2014; Döş, 2016; Gelbal ve Kelecioğlu, 2007; Kaya, Karaçam, Eş ve Tuncel, 2013; Kuran ve Kanatlı, 2009; Metin, 2013; Metin ve Özmen, 2010; Nazlıççek ve Akarsu, 2008; Özdemir, 2010; Sağlam-Arslan, Devocioğlu-Kaymakçı ve Arslan, 2009; Şenel Çoruhlu, Er Nas ve Çepni, 2009; Yaşar ve Sözbilir, 2013).

Yapılan araştırmalarda Türk Eğitim sisteminde alternatif ölçme ve değerlendirme yaklaşımlarından yaygınlıkla performans görevinin (Altun ve Gelbal, 2014; Anıl ve Acar, 2008; Arı, 2010; Dağhan ve Akkoyunlu, 2014; Çiftçi, 2010; Kaya, Karaçam, Eş ve Tuncel, 2013; Metin ve Özmen, 2010; Özdemir, 2010; Şenel Çoruhlu, Er Nas ve Çepni, 2009; Yiğit ve Kırımlı, 2015) öğretmenler tarafından benimsenmiş olduğu ve kullanıldığı görülmektedir. Bunun başlıca nedenlerinden ilki Türkiye’de Millî Eğitim Bakanlığına bağlı resmî ve özel örgün ortaöğretim kurumlarında eğitim, öğretim, yönetim ve işleyişe ilişkin usul ve esaslarını düzenlemeyi kapsayan Millî Eğitim Bakanlığı (MEB) Ortaöğretim Kurumları Yönetmeliğinin (MEB, 2013) 50. Maddesinde yer alan hüküm olduğu söylenebilir. Nitekim 2017 yılında yapılan yeni düzenleme ile (MADDE 50-Değişik: T.C. Resmî Gazete, 2017) “Öğrenciler okulların özelliklerine göre sınavların dışında proje ve performans çalışması ile topluma hizmet etkinliklerine yönelik seminer, konferans ve benzeri çalışmalar yaparlar. Öğrenciler, her dönemde tüm derslerden en az bir performans çalışması, her ders yılında en az bir dersten proje hazırlama görevini yerine getirirler” demektedir. MEB’in ve ilgili mevzuatın öngörmüş olduğu bu düzenleme ile performans görevinin uygulamada sıklıkla ve yaygınlıkla kullanılmasının bir nedeni olarak değerlendirmek mümkündür. Bir diğer neden ise diğer alternatif ölçme ve değerlendirme yaklaşımlarına göre daha kolay uygulanabilir olduğu görüşüdür. Ancak eğitim ve öğretim sürecinde öğrenci performans görevi değerlendirmesi yaklaşımında beklenen işlevsellik, öğrenci performans görevi yaklaşımına yüklenen kavramsal içeriğin uygulamasıyla olasıdır. Bu da öğrenci performans görevi değerlendirmesinde ölçme ve değerlendirmenin nasıl gerçekleştirilmesinin olası olabileceği sorusuna cevap verilmesiyle mümkün olabilir. Ancak performans görevi değerlendirme yaklaşımının uygulamada karşılaşılan bazı ölçme ve değerlendirme sorun ve kısıtlılıkları bulunmaktadır. Nitekim literatürde yapılan araştırmalar incelendiğinde performans ve proje olarak verilecek konuların seçimi, öğrenci seviyesine uygunluk, bir eğitim öğretim döneminde verilecek performans görev sayısı, ödevlerin hazırlanış biçimi, değerlendirme yöntemleri, değerlendirme kriterlerinin seçimi gibi performans görevinin değerlendirilmesi uygulamasında ciddi sayılabilecek düzeyde sorunların olduğu görülmektedir (Arı, 2010; Çiftçi, 2010; Kaya, Karaçam, Eş ve Tuncel, 2013; Metin ve Özmen, 2010). Bu konuda yapılan araştırmalar ayrıntılı olarak incelendiğinde öğretmenlerin alternatif ölçme ve değerlendirme araçları hakkında - performans olarak verilecek konuların nasıl seçileceği, hazırlanacağı, değerlendirme kriterlerinin nasıl belirleneceği ve değerlendirileceği- yeterli bilgi birikimine sahip olmamaları, zaman ve uygulama esnasında sınıf kontrolünde problem yaşamaları, öğretmen boyutundan kaynaklanan sorunlardır. Öğrenciler arasındaki bireysel farklılıklar, performans ve proje gibi ödevlere karşı olumlu veya olumsuz tutumlar, merkezi seviye belirleme ve seçme sınavlarında gösterecekleri başarıyı daha çok önem vermeleri, öğrencilerin performans görevlerini hazırlamaları, sunmaları ve sergilemelerinde sorunlara yol açmaktadır. Ailenin ekonomik durumu, ailenin öğrenciye karşı ilgisi, ailenin proje ve performans görevlerine ilişkin olumsuz tutuma sahip olması, çocuklarının proje ve performans görevleri hakkında hiçbir bilgiye sahip olmama, öğrencilerin performans veya proje görevlerini yerine getirmeleri sürecinde anne babanın ilgisiz olması ya da rehberlik yapamaması, bazı velilerin tam tersine aşırı ilgi göstermeleri, öğrencinin yerine performans görevini tamamen ailelerinin hazırlamaları ise aileden kaynaklanan sorunlardır. Genel olarak zaman, kaynak, malzeme, okulun fiziki kapasitesi, sınıf mevcudu, örnek uygulamaların yetersizliği, uygulayıcıların bilgi eksikliği gibi sorunlar alternatif ölçme ve değerlendirme yöntemlerinin uygulama boyutunda karşılaşılan problemlerdir (Anıl ve Acar, 2008; Arı, 2010; Çiftçi, 2010; Dağhan ve Akkoyunlu, 2014; Kaya, Karaçam, Eş ve Tuncel, 2013; Metin ve Özmen, 2010; Özdemir, 2010; Şenel Çoruhlu, Er Nas ve Çepni, 2009).

Yukarıda ifade edilen hususlar performans görevi değerlendirmesinde öğrencinin başarısının nesnel ve güvenilir bir yaklaşımla değerlendirilmesinin istenilen düzeyde olmadığını göstermektedir. Literatür incelemesinde öğrenci performans görevinin değerlendirmesine ilişkin sistematik bir ölçme ve değerlendirme yaklaşımına rastlanılmamıştır. Öğretmenlerin performans ödevi değerlendirmesinde karşılaştığı sorunların minimize edilmesi performans görevi değerlendirmesinin ölçme ve değerlendirme çerçevesinin geliştirilmesiyle mümkün olabilir. Performans görevi değerlendirmesinin kavramsal çerçevesi ölçme ve değerlendirme uygulamasının bütünlük kantitatif bir yaklaşımı gerektirmektedir. Bununla birlikte performans görevi değerlendirmesinde uygulamada sorunlara neden olan ve cevap bulması gereken bazı ayrıntı hususlar da bulunmaktadır. Bunlardan ilki performans

görevinin değerlendirilmesinde belirlenmiş olan kriterlerin olması gerekmektedir. Çünkü Özçelik (1992)'in belirtmiş olduğu gibi değerlendirmede doğru bir sonuca ulaşılabilmesi ölçütün değerlendirme için uygunluğuyla olasıdır. Her ölçüt, ölçme sonucunun değişik bir amaçla işe yarama koşulunu belirler. Belli bir amaçla değerlendirme yapabilmek için o amaçla kullanılacak, yerinde bir ölçüt aranmalıdır (Özçelik, 1992). Nitekim yapılan bir araştırmada (Metin ve Özmen, 2010) öğretmenlerin öğrencilere verilen performans görevlerinin hangi kriterlere göre değerlendirilmesi gerektiği hususunun önemli olduğunu ifade etmişlerdir. Öğrenci performans görevi değerlendirilmesinde karşılaşılan bir diğer husus muhtemelen birden fazla değerlendirme kriterinin olması durumunda değerlendirmenin nasıl yapılabileceği sorunudur. Bu hususun dikkate alınması öğrenci performans görevinin nesnel değerlendirilmesine olanak sağlaması açısından önemli bulunmaktadır. Buna ilaveten öğrenci performans görevinin değerlendirilmesindeki kriterlerin çok sayıda olması ve muhtemelen performans görevinin kapsamındaki bilim alanı ve konulara göre çeşitlilik gösterebilecektir. Örneğin lise eğitimi düzeyinde verilen matematik dersinin kapsamında hazırlanan bir performans görevi değerlendirmesi ile kimya dersi kapsamında hazırlanmış olan bir performans görevi değerlendirmesinin kriterleri nitelik ve nicelik olarak farklılık gösterebilecektir. Kaldı ki bu tür bir durumla aynı bilim alanında olmakla birlikte konuların içeriklerinin farklılaşmasıyla da karşılaşılabilmektedir.

Örneğin kimya dersinde kimyanın uğraş alanları konusunda hazırlanan bir performans görevi değerlendirmesi ile atom ve periyodik sistem konusunda hazırlanan performans görevi değerlendirilmesine yönelik kriterler sayıca ve nitelik olarak aynı olmayabilecektir. Bütünleşik sistemlerin değerlendirilmesinde karşılaşılan bir özellik olmasından kaynaklanan (Chen, Hsieh, Do, 2015; Lai ve Choi, 2015; Lupo, 2013; Wu, Chen, Chen ve Zhuo, 2012) diğer bir boyut kriterlerin nispi önemleridir. Bir diğer anlatımla performans görevinin değerlendirilmesinde dikkate alınması gereken kriterlerin ölçme ve değerlendirme işleminde nispi önem seviyelerinin aynı olmayabileceğidir. Örneğin bir öğrenci performans görevi değerlendirmesinde temel alınan "n" sayıdaki kriterden biri olan "n1"nin değerlendirme işlemindeki nispi önemi "n2" ya da "nn" kriterlerinden farklı olabilir. Dolayısıyla kriterlerin olası önem farklılıklarını dikkate alan ağırlıklar temelinde öğrenci performans görevi değerlendirmesinin yapılması gerekmektedir. Uygulamada karşılaşılan bir diğer konu ise belirlenmiş olan kriterlere göre öğrenci performans görevinin nasıl ölçüleceğidir. Örneğin kimya dersinde hazırlanmış olan bir performans ödevinde çalışma konusu yapılmış olan maddenin fiziksel hâlleri konusundaki performans çalışmasını kriterler temelinde öğretmenin hangi ölçme aracı ile nasıl ölçmesi gerektiğidir. Yukarıda ifade edilen hususlar öğrenci performans görevinin değerlendirilmesinin sistematik bir yaklaşımla ölçülmesini gerektirmektedir. Yukarıda verilen bilgiler ışığında bu çalışmanın başlıca amacı çok ölçütlü karar verme teknikleri ile öğrenci performans görevinin değerlendirilmesidir.

Bu çalışmanın giriş kısmında ifade edilmiş olan görüşler bağlamında ikinci kısımda öğrenci performans görevi değerlendirilmesine yönelik önerilen yaklaşımın yönetsel çerçevesi ve çalışmada kullanılan teknikler verilmiştir. Çalışmanın üçüncü kısmında 9. sınıf kimya dersi kapsamında kimya bilimi ünitesindeki konu kazanımları (MEB, 2017) dikkate alınarak belirlenmiş olan elementler konusuna ilişkin öğrenci performans görevi değerlendirilmesi bulguları verilmiştir. Dördüncü kısımda ise çalışmanın sonuçları değerlendirilmiş ve düşünülen öneriler verilmiştir.

Yöntem

Çalışmanın başlıca iki aşaması bulunmaktadır. Birinci aşamada öğrenci performans görevinin değerlendirilmesinde temel alınan kriterlerin belirlenmesi ve seçimi yapılmıştır. İkinci aşamada ise belirlenen kriterler temelinde öğrenci performans görevinin ölçülmesi ve değerlendirilmesi verilmiştir.

Bu çalışmada yukarıda ifade edilen yöntemin uygulaması 2017-2018 eğitim ve öğretim yılında 9. sınıf kimya dersinde elementler konusunda verilmiş olan öğrenci performans görevi değerlendirmesini kapsamaktadır. Çalışmada araştırmanın amacına uygun olarak iki şekilde veri sağlanmıştır. Birincisi, çalışmanın verileri ilgili literatürde (Dağdeviren, Yavuz ve Kılınç, 2009; Yüksel ve Geban, 2015; Shyur, 2006) yer alan çok ölçütlü karar verme çalışmalarında olduğu gibi, araştırmada kullanılan tekniklerin özelliklerine uygun bir şekilde uzman görüşünden elde edilmiştir. Bu amaçla çalışmada oluşturulan uzman grup, birinci yazar ile Ankara'da bir lisede görev yapan uzman kimya öğretmeni ve kimya öğretmeni kariyer unvanlarında olan kişilerden oluşmuştur. Çalışmanın birinci yazarı tarafından uzman grubun diğer üyelerine çalışmanın amacı, öğrenci performans görevi, ikili karşılaştırma yapacakları kriterler ve öğrenme alanlarına göre kriterlerin değerlendirilmesi hakkında bilgi verilmiştir. Uzman grup aşağıda belirtilen ikili karşılaştırmaları ve değerlendirmeleri ortak görüşe göre yapmıştır:

- ✓ Öğrenme alanlarına ilişkin ikili karşılaştırmalar.
- ✓ Performans görevi değerlendirme kriterlerinin öğrenme alanlarına göre değerlendirilmesi.
- ✓ Performans görevinde temel olan kriterlerin ana gruplarının ikili karşılaştırılması.

- ✓ Performans görevi değerlendirme kriterlerinin ikili karşılaştırılması.

Uygulamanın birinci aşamasında öncelikle uzman grubun görüşüne göre değerlendirmeler yapılmıştır. Araştırmada oluşturulan uzman grup öğrenme alanlarına ilişkin ikili karşılaştırmalarını, performans görevinde temel olan kriterlerin ana gruplarının ikili karşılaştırmalarını ve performans görevi değerlendirme kriterlerinin ikili karşılaştırmalarını Saaty (1980) tarafından geliştirilen (Tablo 1) ölçekle değerlendirmişlerdir.

Tablo 1. İkili Karşılaştırmalarda Önem Düzeyleri

a_{ij}	Tanım	Açıklama
1	Eşit önem	İki etkinlik eşit derecede amaca katkıda bulunmakta
3	Zayıf önem	Etkinlik değerine kıyaslandığında nispeten tercih edilir
5	Güçlü önem	Etkinlik değerine kıyaslandığında güçlü tercih edilir
7	Çok güçlü ya da kanıtlanmış önem	Etkinlik değerine kıyaslandığında çok güçlü tercih edilir
9	Mutlak önem	Etkinliğin değerine tercih edilmesindeki kanıtın çok yüksek güvenilirlik derecesi vardır
2,4,6,8	Ara değerler	İhtiyaç olduğunda kullanılır

Çalışmanın ikinci aşamasında ise öğrenci performans görevinin değerlendirmesi uzman grupta yer alan 9. sınıf kimya dersini veren öğretmen tarafından yapılmıştır. Bulguların birinci ve ikinci aşamasındaki değerlendirmelere ilişkin verilerin analizi Expert Choice (2000) ve Microsoft Excel (2016) programları ile yapılmıştır.

Araştırmada Kullanılan Çok Ölçütlü Karar Verme Teknikleri

Öğrenci performans görevi değerlendirmesini konu edinen bu çalışmanın uygulamasında çok ölçütlü karar verme tekniklerinden TOPSIS (Technique for Order Performance by Similarity to İdeal Solution) ve AHP (Analitik Hiyerarşi Prosesi) kullanılmıştır.

AHP Tekniği

Saaty (1980)'nin geliştirmiş olduğu AHP tekniği literatürde çok ölçütlü karar verme (Gupta ve Chaudary, 2017; Lupo, 2013; Wu, Chen, Chen ve Zhuo, 2012; Zyoud ve Fuchs-Hanusch, 2017) problemlerinde sıklıkla kullanılan bir yaklaşımdır. AHP tekniğinin başlıca özelliklerinden ilki analiz konusu olan problemin öğelerden oluşan bir bütün niteliğinde olmasıdır. Bu özelliğinden dolayı AHP tekniğinde öncelikle araştırma problemini oluşturan öğeler, amaç fonksiyonuna göre hiyerarşik olarak düzenlenmektedir. Bu tür bir ayrıştırma araştırma problemini oluşturan bütün öğelerin modelde olmasını sağlamaktadır. Böylece, oluşturulan modelde yer alan öğelerin hiyerarşik seviyelerine ve gruplandırılmasına göre ikili karşılaştırma matrisleri oluşturulmaktadır (Saaty, 1980; Saaty, 1986). Buna göre AHP tekniğinde i. unsuru ile j. unsurun ikili karşılaştırılması a_{ij} ile belirtilmektedir. Buna karşılık modeldeki j. unsur ile i. unsurun karşılaştırılması a_{ji} ile gösterilmektedir. Diğer bir ifadeyle AHP'de a_{ji} unsurunun değeri $a_{ji} = 1/a_{ij}$ eşitliğinin değerine karşılıktır. Genellikle AHP modelindeki unsurların ikili karşılaştırmalarında Saaty'nin (1980) önerdiği ölçek (Tablo 1) kullanılmaktadır. AHP tekniğinin bir diğer özelliği ikili karşılaştırma matrislerinin $(n - 1)/2$ kadar ikili karşılaştırmaların yapıldığı $n \times n$ boyutunda kare matrislerdir.

AHP'de yapılan ikili karşılaştırmalar neticesinde $n \times n$ boyutlu kare matristeki ifadesi şu şekildedir (Saaty, 1980):

$$A = \begin{bmatrix} 1 & a_{12} & \dots & a_{1n} \\ 1/a_{12} & 1 & \dots & a_{2n} \\ \dots & \dots & 1 & \dots \\ 1/a_{1n} & 1/a_{2n} & \dots & 1 \end{bmatrix}$$

Öncelik vektörü $(A - \lambda_{max} I)W = 0$ eşitliği ile belirlenmektedir. Eşitlikte A ikili karşılaştırma matrisini, W özvektör ve λ_{max} A matrisinin en büyük özdeğerini belirtmektedir. AHP'nin diğer bir özelliği ve üstünlüğü değerlendirmelere ilişkin ikili karşılaştırmaların ne düzeyde tutarsız olabileceğinin belirlenebilmesidir (Saaty, 1980; Saaty,

1991). Tutarlılık Oranının ($T.O$) belirlenmesi için Tutarlılık İndeksi ($T.I$) ve Rassal İndeksin ($R.I$) hesaplanması gerekmektedir. Tutarlılık İndeksi $T.I = (\lambda_{max} - n)/(n-1)$ eşitliğiyle hesaplanmaktadır. Eşitlikte yer alan n matrisin boyutunu ifade etmektedir. Rassal indeks değeri ise matrisin boyutuna (Tablo 2) göre belirlenmektedir. Buna göre ikili karşılaştırma matrisinin tutarsızlık oranı $T.O=T.I/R.I$ eşitliğiyle hesaplanmaktadır. Matrisin boyutuna göre farklılık göstermekle birlikte ikili karşılaştırmaların tutarlılığının 0.10'dan küçük olması mümkündür. AHP'de tutarsızlık oranının 0.10'dan büyük olması halinde ikili karşılaştırmalar tekrar yapılmaktadır. Modeldeki öğelere ilişkin ikili karşılaştırmaların yapılması, önceliklerin ve tutarlılık oranlarının hesaplanması modeldeki bütün hiyerarşik seviyelerdeki gruplandırmalar için yapılmaktadır.

Tablo 2. Rasgele İndeks Değerleri (RI) (Saaty, 1980)

(n)	1	2	3	4	5	6	7	8	9
(RI)	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45

TOPSIS Tekniği

Çalışmada kullanılan diğer çok ölçütlü karar verme tekniği TOPSIS ise çok sayıda olan alternatifin birden fazla kriter bağlamında analizine ve değerlendirilmesine imkân vermektedir. Çalışmada TOPSIS tekniğinin öğrenci performans görevi değerlendirmesinde kullanılacak kriterlerin belirlenmesinde en uygun olduğu düşünülen bir yaklaşım olduğu için bu çalışmada tercih edilmiştir. Çünkü TOPSIS tekniğinin temel özelliklerinden biri birden fazla kriter bağlamında seçeneklerin sıralanmasına imkân vermesidir. Bu çalışmada da öğrenci performans görevinin değerlendirilmesinde kullanılacak kriterlerin seçiminin birden fazla boyut altında olabileceği düşüncesiyle TOPSIS tekniği kullanılmıştır. Çok ölçütlü karar verme tekniklerinden TOPSIS ilk kez Hwang ve Yoon (1981) tarafından geliştirilmiştir. TOPSIS tekniği çeşitli çok ölçütlü problemlerin çözümlenmeye çalışıldığı (Dağdeviren, Yavuz ve Kılınç, 2009; Ersöz, Kabak ve Yılmaz, 2011; Zyoud ve Fuchs-Hanusch, 2017) araştırmalarda kullanılmıştır. TOPSIS tekniğinin başlıca adımları şu şekildedir (Shyur, 2006):

Adım 1: Sıralama için karar matrisinin oluşturulması. Karar matrisinin yapısı aşağıda gösterildiği gibidir. A_j muhtemel alternatifleri göstermektedir; $i = 1, \dots, m$; F_i ise ilgili i . özellik ya da kriteri, $j = 1, \dots, n$; f_{ij} ise bir değer olmak üzere her bir alternatif A_i ile her bir F_j kriterinin karşılaştırılması durumunda.

$$D = \begin{matrix} & F_1 & F_2 & \dots & F_j & \dots & F_n \\ A_1 & \left[\begin{array}{cccccc} f_{11} & f_{12} & \dots & f_{1j} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2j} & \dots & f_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ f_{i1} & f_{i2} & \dots & f_{ij} & \dots & f_{in} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ f_{j1} & f_{j2} & \dots & f_{jj} & \dots & f_{jn} \end{array} \right. \end{matrix}$$

Adım 2: standartlaştırılmış karar matrisinin hesaplanması. $R(= [r_{ij}])$.normalleştirilmiş değer r_{ij} şu eşitlikle hesaplanmaktadır: $r_{ij} = \frac{f_{ij}}{\sqrt{\sum_{j=1}^n f_{ij}^2}}$, $j = 1, \dots, n$; $i = 1, \dots, m$.

Adım 3: Ağırlıklandırılmış normalize karar matrisinin hesaplanması. Bu normalleştirilmiş karar matrisi ile ilgili ağırlıkların çarpımıyla hesaplanmaktadır. Ağırlıklandırılmış normalize değerler v_{ij} şu şekilde hesaplanmaktadır: $v_{ij}=w_j r_{ij}$, $j = 1, \dots, n$; $i = 1, \dots, m$, w_j ifadesi j . özellik ya da kriterin ağırlığını göstermektedir.

Adım 4. İdeal ve negatif ideal çözümlerin saptanması. J fayda kriteri ile ilgili olmak üzere ve J' ise maliyet kriterini göstermek üzere şu şekildedir:

$$V^+ = \{v_1^+, \dots, v_n^+\} = \{(\max_i v_{ij} | j \in J), (\min_i v_{ij} | j \in J')\}$$

$$V^- = \{v_1^-, \dots, v_n^-\} = \{(\min_i v_{ij} | j \in J), (\max_i v_{ij} | j \in J')\},$$

Adım 5. *Ayırma ölçülerinin hesaplanması.* Bu m-boyutlu Euclidean uzaklık kullanılarak hesaplanmaktadır. Her bir alternatifi ideal çözümden ayırma (D_i^+) şu şekildedir:

$$D_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2}, i = 1, \dots, m, \text{ benzer şekilde her bir alternatifi negatif ideal çözümden ayırma ise } (D_i^-) \text{ şöyledir: } D_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}, i = 1, \dots, m$$

Adım 6. *ideal çözüme nispi yakınlığın hesaplamak ve tercihlerin sıralamasını belirlemek.*

Hesaplama ve sıralama $C_i = \frac{D_i^-}{D_i^+ + D_i^-}, i = 1, \dots, m,$ eşitlikle gösterilmiştir. C_i indeks değeri 0 ile 1 aralığında değer almaktadır. İndeks değerlerinin büyük olması daha iyi performans alternatifi göstermektedir.

Bulgular

Çalışmanın bu kısmında 9. sınıf kimya dersi kapsamındaki elementler konusu bağlamında öğrenci performans görevinin değerlendirilmesine yönelik bulgular verilmiştir.

Bu çalışmanın ilk aşamasında öğrenci performans görevinin değerlendirilmesinde temel alınan kriterlerin oluşturulması ve seçimi yapılmıştır. Bunun için önce öğrenci performans görevinin değerlendirilmesinde temel alınan kriterler belirlenmiştir. İlgili literatürde kimya dersi ya da içeriğini değerlendirmeye yönelik spesifik kimya konusu bağlamındaki kriterlere rastlanılmamıştır. Ancak Ortaöğretim 9. Sınıf Kimya Dersi Öğretim Programı (MEB, 2007) ve Orta Öğretim 9.Sınıf Kimya Ders Kitabında (Dursun, Gülbay, Çetin ve Tek, 2012; Güntut, Güneş ve Çetin, 2017) bazı genel performans görevi değerlendirme kriterlerinin mevcut olduğu ve Türk Eğitim Sistemi'nde kullanılması için önerilmiş olduğu görülmüştür. Bu çalışmada temel alınan kriterlerin belirlenmesinde bu (Dursun, Gülbay, Çetin ve Tek, 2012; Güntut, Güneş ve Çetin, 2017; MEB, 2007) kaynaklar esas alınmış olmakla birlikte elementler konusunu temsil eden kriterler bu çalışmada geliştirilmiştir (Tablo 3).

Tablo 3. Öğrenci Performans Görevi Değerlendirme Kriterleri

Kod	Kriterler
P01	Tanımlayıcı bilgileri verebilme
P02	Fiziksel özellikleri verebilme
P03	Kimyasal özellikleri verebilme
P04	Periyodik sistemdeki yerini gösterebilme
P05	Keşfine ilişkin bilgileri verebilme
P06	Günlük hayattaki kullanım alanlarını açıklayabilme
P07	Endüstrideki temel kullanım alanlarını verebilme
P08	İnsan sağlığına etkilerini açıklayabilme
P09	Çevre sağlığına etkilerini açıklayabilme
P10	Performans görevi amacını belirleyebilme
P11	Performansa uygun çalışma planı yapabilme
P12	İhtiyaç duyulan kaynakları belirleyebilme
P13	Konuya ilişkin çok fazla ve çeşitli kaynakları kullanabilme
P14	Kullanılan kaynakları konuyla doğrudan ilişkilendirebilme

P15	Bilgilerin güncel olması
P16	Konuya ilişkin kavram, olgu ve prensipleri doğru ve yerinde kullanabilme
P17	Yorumları kaynaklardaki kanıtlarla destekleyebilme
P18	Türkçe dilbilgisi yazım kurallarını kullanabilme
P19	Eleştirel düşünme becerisini gösterebilme
P20	Yaratıcılık yeteneğini kullanabilme
P21	Toplanan bilgileri analiz ederek düzenleyebilme
P22	Performans çalışmasını plana göre gerçekleştirebilme
P23	Türkçeyi doğru ve düzgün konuşabilme
P24	Sunuyu hedefe yönelik materyalle destekleyebilme
P25	Konuyu dinleyicilerin ilgisini çekecek şekilde sunabilme
P26	Sorulara cevap verebilme
P27	Sunuda akıcı bir dil ve beden dilini kullanabilme
P28	Sunum yaparken süreyi etkin kullanabilme.
P29	Sunum sırasında özgüvene sahip olabilme

Tablo 3'teki kriterler iki aşamadan oluşan bir analize tabi tutulmuştur: Birinci aşamada 9. sınıf kimya dersi içeriğinde elementler konusunda hazırlanan bir performans görevini değerlendirmede yeterli sayılabileceği düşünülen 29 kriter belirlenmiştir. İkinci aşamada ise belirlenmiş olan 29 kriterin elementler konusunda hazırlanan bir performans görevini hangi önem derecesinde ölçmeye yönelik olduğu TOPSIS tekniğiyle belirlenmiştir. TOPSIS analizi için öncelikle öğrenci performans görevi değerlendirme kriterlerinin seçiminde esas alınacak boyutların ağırlıkları AHP tekniğiyle hesaplanmıştır. Bu çalışmada kriterlerin performans görevi değerlendirmesindeki önem derecesinin belirlenmesinde bilişsel, duyuşsal ve devinişsel (psikomotor) öğrenme alanları esas alınmıştır. Buna göre oluşturulan ikili karşılaştırma matrisi ve hesaplanan yerel ağırlıklar Tablo 4'te sunulmuştur. Yapılan hesaplamada ikili karşılaştırma matrisinin tutarlılık oranı ($TO = 0.05$) kabul edilebilir bulunmuştur.

Tablo 4. Öğrenme Alanlarının İkili Karşılaştırmaları ve Ağırlıkları

Öğrenme Alan Boyutları	BS	DS	DE	Ağırlıklar
Bilişsel (BS)	1	2	2	0.493
Duyuşsal (DS)		1	2	0.311
Devinişsel (DE)			1	0.196

Çalışmada performans görevini değerlendirmek için gerekli olan kriterler öğrenme alanları (bilişsel, duyuşsal ve devinişsel) boyutlarına göre ağırlıkları belirlenmiş ve sonrasında TOPSIS tekniği ile önem derecelerine göre kriterlerin sayısı azaltılmıştır. Bunun nedeni az sayıda ancak performans görevini değerlendirmede geçerliliği olan kriterlerin kullanılmasını ve böylece performans görevi değerlendirme uygulama sürecinin nesnelleştirilmesi ve kolaylaştırılmasıdır. TOPSIS analizi sonucunda (Tablo 5) önem derecesi dikkate alınarak performans değerlendirmede temel olacak kriterler belirlenmiştir. Performans görevi değerlendirmesinde temel alınan kriterlerin belirlenmesindeki kritik değer olarak aritmetik ortalama ve standart sapma değeri dikkate alınmıştır. Buna göre TOPSIS analizi sonucunda hesaplanmış olan sıralama değerlerinin aritmetik ortalamasının (μ) bir standart sapmanın ($-\sigma$) altındaki değer üzerinde yer alan kriterler seçilmiştir. Böylece kriterlerin seçiminde normal dağılım eğrisi altında kalan alan itibarıyla %84,1'lik aralık içerisinde $[0,50+ (1\sigma)]$ bulunan kriterler yer almıştır. Bunun için öncelikle TOPSIS analizi önem derecelerinin aritmetik ortalaması (3,0267586) ve standart sapması (0,4320363) hesaplanmıştır. Kritik değer ise $(\mu-\sigma)$ 2,594722 olarak hesaplanmıştır. Bu kritik değerden büyük önem derecesine sahip olan kriterler performans değerlendirmesinde temel alınmıştır (Tablo 5).

Tablo 5. TOPSIS Analizine Alınan Kriterlerin Öğrenme Alanlarına Göre Değerlendirilmesi

Kod	Kriterler	Bilişsel	Duyuşsal	Devinişsel	Sıralama Değeri
P01	Tanımlayıcı bilgileri verebilme	0.493	0.311	0.196	3.790
P02	Fiziksel özellikleri verebilme	5	1	2	3.168
P03	Kimyasal özellikleri verebilme	5	1	2	3.168

P04	Periyodik sistemdeki yerini gösterebilme	5	3	2	3.790
P05	Keşfine ilişkin bilgileri verebilme	4	3	3	3.493
P06	Günlük hayattaki kullanım alanlarını açıklayabilme	4	3	3	3.493
P07	Endüstrideki temel kullanım alanlarını verebilme	4	3	2	3.297
P08	İnsan sağlığına etkilerini açıklayabilme	4	3	3	3.493
P09	Çevre sağlığına etkilerini açıklayabilme	4	3	3	3.493
P10	Performans görevi amacını belirleyebilme	4	2	2	2.986
P11	Performansa uygun çalışma planı yapabilme	4	2	2	2.986
P12	İhtiyaç duyulan kaynakları belirleyebilme	4	2	3	3.182
P13	Konuya ilişkin çok fazla ve çeşitli kaynakları kullanabilme	4	2	2	2.986
P14	Kullanılan kaynakları konuyla doğrudan ilişkilendirebilme	3	2	2	2.493
P15	Bilgilerin güncel olması	3	2	2	2.493
P16	Konuya ilişkin kavram, olgu ve prensipleri doğru ve yerinde kullanabilme	4	3	2	3.297
P17	Yorumları kaynaklardaki kanıtlarla destekleyebilme	3	3	2	2.804
P18	Türkçe dilbilgisi yazım kurallarını kullanabilme	3	3	3	3.000
P19	Eleştirel düşünme becerisini gösterebilme	3	3	3	3.000
P20	Yaratıcılık yeteneğini kullanabilme	3	3	3	3.000
P21	Toplanan bilgileri analiz ederek düzenleyebilme	3	2	3	2.689
P22	Performans çalışmasını plana göre gerçekleştirebilme	3	3	3	3.000
P23	Türkçeyi doğru ve düzgün konuşabilme	3	2	3	2.689
P24	Sunuyu hedefe yönelik materyalle destekleyebilme	3	3	3	3.000
P25	Konuyu dinleyicilerin ilgisini çekecek şekilde sunabilme	3	2	2	2.493
P26	Sorulara cevap verebilme	3	3	3	3.000
P27	Sunuda akıcı bir dil ve beden dilini kullanabilme	2	1	2	1.689
P28	Sunum yaparken süreyi etkin kullanabilme.	3	3	3	3.000
P29	Sunum sırasında özgüvene sahip olabilme	3	3	2	2.804

Tablo 5'te görüldüğü üzere TOPSIS analizi neticesinde sıralama değerine göre belirlenen 25 kriter içeriklerine göre dört ana kriter altında sınıflandırılmıştır (Tablo 6) ve bu kriterler yeniden kodlanmıştır. Tablo 6'da yapılan sınıflandırmaya göre AHP tekniği ile ana ve alt kriterlerin ikili karşılaştırmaları yapılmıştır (Tablo 6, Tablo 7). İkili karşılaştırmalardan sonra Expert Choice (2000) programı ile öncelikle ana grupların ve sonra kriterlerin ağırlıkları hesaplanmıştır. Ana gruplara ilişkin ikili karşılaştırma matrisinin (Tablo 6) tutarlılık oranının (T.O) 0.09 olduğu ve bu oranın ikili karşılaştırmaların tutarlı olduğunu göstermiştir.

Tablo 6. Performans Görevinde Temel Olan Kriterlerin Ana Gruplarının İkili Karşılaştırması

Gruplar	E	H	I	S
Elementin özellikleri (E)	1	5	3	7
Hazırlık süreci (H)		1	1/3	5
İçerik (I)			1	5
Sunu yapma (S)				1

Tablo 7. Performans Görevi Değerlendirme Kriterlerinin İkili Karşılaştırması

Kriterler										
ELEMENT ÖZELLİKLERİ (E)	E1	E2	E3	E4	E5	E6	E7	E8	E9	
E1 Tanımlayıcı bilgileri verebilme (E1)	1	3	3	3	5	5	5	5	5	
E2 Fiziksel özellikleri verebilme (E2)		1	1	1/3	3	3	3	3	3	
E3 Kimyasal özellikleri verebilme (E3)			1	1/3	3	3	3	3	3	
E4 Periyodik sistemdeki yerini gösterebilme (E4)				1	3	3	3	3	3	
E5 Keşfine ilişkin bilgiyi verebilme (E5)					1	1/3	1/3	1/3	1/3	
E6 Günlük hayattaki kullanım alanlarını açıklayabilme (E6)						1	3	1	1	

E7	Endüstrideki temel kullanım alanlarını verebilme (E7)	1	1/3	1/3				
E8	İnsan sağlığına etkilerini açıklayabilme (E8)		1	3				
E9	Çevre sağlığına etkilerini açıklayabilme (E9)						1	
HAZIRLIK SÜRECİ (H)		H1	H2	H3	H4			
H1	Performans görevi amacını belirleyebilme (H1)	1	3	3	3			
H2	Performansa uygun çalışma planını yapabilme (H2)		1	1/3	1/3			
H3	İhtiyaç duyulan kaynakları belirleyebilme (H3)			1	1			
H4	Konuya ilişkin çok fazla ve çeşitli kaynakları kullanabilme (H4)				1			
İÇERİK (I)		I1	I2	I3	I4	I5	I6	I7
I1	Konuya ilişkin kavram, olgu ve prensipleri kullanabilme (I1)	1	1	3	3	3	1	3
I2	Yorumların kaynaklardaki kanıtlarla destekleyebilme (I2)		1	3	3	3	1	3
I3	Türkçe dilbilgisi yazım kurallarını kullanabilme (I3)			1	1/3	1/3	1/3	1/3
I4	Eleştirel düşünme becerisini gösterebilme (I4)				1	1	1	3
I5	Yaratıcılık yeteneğini kullanabilme (I5)					1	1	3
I6	Toplanan bilgileri analiz ederek düzenleyebilme (I6)						1	3
I7	Performans çalışmasını plana göre gerçekleştirebilme (I7)							1
SUNU YAPMA (S)		S1	S2	S3	S4	S5		
S1	Türkçeyi doğru ve düzgün konuşabilme (S1)	1	1/3	1/3	3	3		
S2	Sunuyu hedefe yönelik materyalle destekleyebilme (S2)		1	1	3	3		
S3	Sorulara cevap verebilme (S3)			1	3	3		
S4	Sunum yaparken süreyi etkin kullanabilme (S4)				1	1		
S5	Sunum sırasındaki özgüvene sahip olabilme (S5)					1		

Çalışmanın ikinci aşamasında öğrenci performans görevinin değerlendirilmesi kriterler ve hesaplanan ağırlıklar temelinde yapılmıştır. Çalışmada öğrencinin performans görevi her bir kritere göre Tablo 8’de verilen ölçek ile değerlendirilmiştir. Altı dereceden oluşan bu ölçek literatürde çok ölçütlü karar verme problemlerinde kullanılan (Yüksel ve Dağdeviren, 2006; Yüksel, 2011) bir ölçektir. Değerlendirme ölçeğinin değerleri 0.0 ile 1.0 arasında yer almaktadır.

Tablo 8. Değerlendirme Ölçeği

Değerlendirme Düzeyleri	Düzye Değeri
Çok iyi (CI)	1.0
İyi (IY)	0.8
Orta (OR)	0.6
Olumsuz (OZ)	0.4
Çok Olumsuz (CO)	0.2
Değerlendirmeye Giremez (DG)	0.0

Çalışmada Tablo 8’deki ölçek ile belirlenen kriterlere göre bir öğrencinin elementler konusunda hazırlamış olduğu öğrenci performans görevinin değerlendirilmesi yapılmıştır. Tablo 9’un ikinci sütununda verilen her bir kriter temelinde öğrencinin performans görevinin değerlendirilmesi Tablo 8’in düzeylerine göre yapılmıştır. Tablo 9’un beşinci sütununda değerlendirmenin düzeyi ve altıncı sütunda ise düzeye karşılık olan değer verilmiştir. Tablo 9’un son sütununda ise öğrencinin her bir kritere göre performans düzeyi verilmiştir. Bu değer her bir kriterin genel ağırlık değeri ile altıncı sütundaki değer çarpımından oluşan bir değerdir. Tablo 9’un son satırında ise öğrencinin bütün kriterler temelinde değerlendirilmesi durumunda almış olduğu puanı vermektedir. Bu çalışmanın uygulaması kapsamında değerlendirilen öğrencinin performans görevi puanının 81.2 (100) olduğu saptanmıştır.

Tablo 9. Kriter Ağırlıkları ve Öğrenci Performans Görevi Değerlendirme Sonuçları

Gruplar ve Ağırlıklar	Kriterler	Yerel Ağırlıklar	Genel Ağırlıklar (ga)	Mevcut Durum	Değer (d)	Performans Düzeyi (ga×d)
Element Özellikleri (E) 0.556	E1	0.299	0.166	CI	1.0	0.166
	E2	0.123	0.069	CI	1.0	0.069
	E3	0.123	0.069	CI	1.0	0.069
	E4	0.186	0.103	CI	1.0	0.103
	E5	0.032	0.018	IY	0.8	0.014
	E6	0.063	0.035	OR	0.6	0.021
	E7	0.041	0.023	IY	0.8	0.018
	E8	0.075	0.041	OR	0.6	0.025
	E9	0.058	0.032	OR	0.6	0.019
Hazırlık Süreci (H) 0.136	H1	0.487	0.066	IY	0.8	0.053
	H2	0.096	0.013	IY	0.8	0.010
	H3	0.208	0.028	IY	0.8	0.022
	H4	0.208	0.028	OR	0.6	0.017
İçerik (I) 0.259	I1	0.233	0.060	IY	0.8	0.048
	I2	0.233	0.060	OR	0.6	0.036
	I3	0.049	0.013	IY	0.8	0.010
	I4	0.124	0.032	OR	0.6	0.019
	I5	0.124	0.032	OR	0.6	0.019
	I6	0.166	0.043	OR	0.6	0.026
	I7	0.069	0.018	IY	0.8	0.014
Sunu(S) 0.049	S1	0.175	0.009	IY	0.8	0.007
	S2	0.326	0.016	OR	0.6	0.009
	S3	0.326	0.016	IY	0.8	0.013
	S4	0.086	0.004	IY	0.8	0.003
	S5	0.086	0.004	OR	0.6	0.002
Öğrenci Performans Görevi Düzeyi						0.812

Tartışma ve Sonuç

Bu çalışmada kuramsal temelleri yapısalıcı bilgi kuramına dayanan alternatif ölçme ve değerlendirme yöntemlerinden biri olan ve Türkiye'deki ortaöğretim düzeyinde kullanılan öğrenci performans görevi değerlendirmesinde karşılaşılan bazı sorunlar çok ölçütlü karar verme yaklaşımlarıyla çözümlenmeye çalışılmıştır. Çalışmanın sonuçlarına yöntemsel açıdan bakıldığında, performans kavramının genellikle tek bir bileşenden oluşmayan doğası gereği ve bu çalışmanın çözmeyi hedeflediği sorunların yapısından dolayı bu çalışmada bütünleşik değerlendirmelere imkân veren çok ölçütlü karar verme yaklaşımları kullanılmıştır. Nitekim literatürde de performans olgusunun araştırma konusu yapılmış olduğu (Lai ve Choi, 2015) ya da birden fazla kriterin yer aldığı, kriter seçim probleminin olduğu (Chen, Hsieh, Do, 2015; Gupta ve Chaudary, 2017) çalışmaların çok ölçütlü tekniklerle problemleri çözmeye çalıştıkları görülmektedir. Bu çalışmada kullanılan tekniklerin tercih edilmesinin esas gerekçesi yukarıda ifade edilen düşünce bağlamında, AHP ve TOPSIS'in bütüncül bir yapıda olan ve çok sayıda unsuru birlikte değerlendirebilen analitik bir uygulamaya sahip olmasıdır. Nitekim bu çalışmada kullanılan tekniklerin araştırma amacına uygunluğunun literatürde yer alan araştırmalarda (Chen, Hsieh, Do, 2015; Gupta ve Chaudary, 2017; Lupu, 2013; Wu, Chen, Chen ve Zhuo, 2012; Zyoude ve Fuchs-Hanusch, 2017) kullanılan tekniklerin sonuçlarıyla uyumlu bulunmaktadır.

Literatürde öğrenci performans görevinin değerlendirilmesinin uygulamasını sistematik ve çok ölçütlü bir yaklaşımla çalışma konusu yapan araştırmalara henüz rastlanılmadığı için bu çalışmanın sonuçları kendi içerisinde tartışılmıştır. Öğrenci performans görevinin değerlendirilmesinin sistematik bir yaklaşımla değerlendirilebilmesi

için öğrenci performans görevinin değerlendirilmesinde temel alınacak kriterlerin belirlenmesi ölçme ve değerlendirilmenin geçerliliği açısından önemli bulunmaktadır. Bu nedenle bu çalışmada önerilen yöntemle yapılan uygulamanın sonuçlarından ilki performans görevinin değerlendirilmesinde temel alınan kriterler belirlenmesidir. Değerlendirme kriterlerinin belirlenmesi ölçme ve değerlendirilmenin hangi çerçevede yapılacağını göstermektedir. Böylece ölçülmek ya da değerlendirilmek istenenin geçerli ve güvenilir bir düzlemde ölçülmesi sağlanmıştır. Aksi halde kriterleri belli olmayan bir öğrenci performans görevinin değerlendirmesinin anlamlı olması beklenemez.

Çalışmanın bir diğer sonucu öğrenci performans görevinin değerlendirmesinde kullanmak amacıyla belirlenen kriterlerin sayısının TOPSIS tekniğiyle azaltılmış olmasıdır. Bu işlemin sonucu öğretmenlerin daha az sayıda kriterle çalışmalarına imkân vererek öğrenci performans görevi değerlendirmelerinin daha kolay ve zaman sorunu ile karşılaşmadan uygulamaya fırsat sağlayabilir. Nitekim yapılan araştırmalarda (Anıl ve Acar, 2008; Dağhan ve Akkoynlu, 2014; Gelbal ve Kelecioğlu, 2007; Kaya vd, 2013;) alternatif ölçme ve değerlendirme yöntemlerinin kullanılmamasının nedenleri arasında uygulama sürecinin güçlüğü ve zaman gerektirici özelliğinin olduğu ifade edilmiştir.

Bu çalışmanın sonuçlarının ilgili literatür açısından bir diğer önemi öğrenci performans görevi değerlendirilmesinin yargısal veya sezgisel bir yaklaşımla değil çok ölçütlü karar verme tekniklerinden biri olan TOPSIS ile nesnel bir yaklaşımla kriterlerin belirlenebilmiş olmasıdır. Bu tür bir seçim kriterlerin birden fazla unsur temelinde irdelemesiyle ölçme ve değerlendirmede temsil gücü yüksek olan kriterlerin belirlenmesine imkân vermektedir.

Vurgulanması gereken bir diğer husus geçerli bir ölçmeden söz edebilmek için ölçme ve değerlendirilmenin konunun ya da bütünsel olarak öğretim programının kazanımlarına yönelik alanları dikkate alan bir ölçme ve değerlendirme yaklaşımı olması gerekmektedir. Bu düşünceden hareketle bu çalışmada kriter seçimi öğrenci kazanımlarının yönelik olduğu bilişsel, duyuşsal ve psikomotor alan hedefleri temelinde yapılmıştır. Böylece öğrenme alanlarına etkisi olmayan ya da düşük katkısı olan kriterler ölçme ve değerlendirmede kullanılmamıştır. Bu ise ölçme ve değerlendirilmenin geçerliliğini kuvvetlendiren bir sonuçtur.

Bu çalışmanın diğer bir özelliği, değerlendirme kriterlerinin nispi ağırlıklarının dikkate alınarak AHP ile belirlenmiş olmasıdır. Kriterlerin nispi ağırlıklarının belirlenmesi öğrenci performans görevinin değerlendirilmesinin geçerliliği açısından önemli bir işlevi bulunmaktadır. Çünkü herhangi bir öğrenci performans görevinin değerlendirilmesinde temel alınacak kriterlerin tamamının aynı önem düzeyinde ya da ağırlıkta olmasını beklemek ve söylemek güçtür. Doğası gereği kriterin içeriğine bağlı olarak kriterlerin nispi ağırlıkları ve önemleri farklı olabilecektir. Bu nedenle bu çalışmada öğrenci performans görevinin değerlendirilmesinde kullanılacak kriterlerin ne düzeyde ağırlıklı oldukları belirlenmiştir. Öte yandan kriterlerin ağırlıklarının belirlenmesindeki uzman görüşünün ne düzeyde tutarlı olduğuna ilişkin hesaplamalar yapılmış ve tutarlılık oranları bulunmuştur. Diğer bir ifadeyle uzman görüşlerinin olası tutarsızlıklarının değerlendirme sürecine yansımaması sağlanmıştır. Böylece öğrenci performans görevinin değerlendirilmesi güvenilir bir yaklaşımla belirlenebilmiştir.

Çalışmanın sonuçlarına göre çalışmanın nihai amacı olan 9. sınıf kimya dersi kapsamındaki elementler konusu bağlamında öğrenci performans görevinin değerlendirilmesi ayrıntılı bir şekilde gerçekleştirilmiş ve öğrenci performans görevinin değerlendirilmesi element özellikleri, hazırlık süreci, içerik ve sunu ana grupları altında toplanan 25 alt kriter temelinde değerlendirilmiştir. Çalışmada öğrencinin performans görevi değerlendirmesi bütünsel olarak değerlendirilmiş ve buna göre öğrencinin başarısı (%81,2) belirlenmiştir. Bununla birlikte öğrencinin performans görevinin her bir ana gruba ve her bir kritere göre nasıl bir performans gösterdiği belirlenebilmiştir. Performans görevinin ayrıntılı değerlendirilmiş olması öğrencinin eksik ve yetersiz yönlerinin belirlenmesine imkân verdiği gibi öğrencinin geliştirilmesi gereken yönlerine ilişkin bilgi de vermektedir. Böylece performans değerlendirilmenin işlevi yalnızca bir sonuç belirleme değil, aynı zamanda eğitimin amaçlarını öğrenciye kazandırmanın bir süreci olarak kullanılmasına olanak sağlayacaktır.

Çalışmanın sonuçları öğrenci performans görevinin ölçme ve değerlendirme işleminin çok boyutlu bir yaklaşım gerektirdiğini ve bu özelliğinin uygulama sürecine yansıtılmasını gerektiğini göstermiştir. Çalışmanın yöntem ve uygulama açısından Türk eğitim sistemindeki öğrenci performans görevi uygulamasına katkı sağlayacağı düşünülmektedir. Çünkü mevcut ortaöğretim sisteminde performans görevi bir zorunlu değerlendirme yöntemi olarak uygulamada yer almış olmasına karşılık öğretmenlerin nasıl bir değerlendirme uygulaması yapacakları konusunda sistematik bir çerçeve bulunmamaktadır. Bu çalışmanın yönetsel boyutu ve uygulama sonuçları performans görevi değerlendirme konusundaki sistematığın geliştirilmesinde öncü ve katkı sağlayabileceği söylenebilir. Örneğin öğretmenlerin kimya dersine yönelik öğrencilerin performans değerlendirme kriterleri bu çalışmada önerilen yaklaşımla belirlenebilir. Öğretmenler belirlenmiş olan kriterler temelinde öğrencilerin performans görevlerini nesnel

bir yaklaşımla belirlemesine olanak sağlayabilir. Ayrıca performans ödevinin değerlendirilmesinde öğretmene yönelik çeşitli paydaşlardan kaynaklardan baskılar da giderilmiş olacaktır.

Bu çalışmanın devamında yapılabilecek çalışmalardan ilki bulanık sayılar (Şen, 2001) yaklaşımıyla ölçme ve değerlendirmenin yapılmasıdır. Bu çalışmada performans değerlendirme kriterlerinin seçiminde, ağırlıkların belirlenmesinde ve kriterler temelinde öğrenci performans görevinin değerlendirilmesinde kesin sayılar kullanılmıştır. Bilindiği gibi kesin sayılar Aristo mantığına dayalı klasik küme kavramına göre tanımlanmış olan gerçek sayılardır (Şen, 2001). Ancak kesin sayıların karar verme ve değerlendirme problemlerinde her zaman durum ve koşulları tanımlaması mümkün olmayabilir. Bu sebeple ölçme ve değerlendirmelerin bulanık bir yaklaşımla yapılması daha uygun bir değerlendirmeye olanak sağlayacağı düşüncesinden dolayı gelecekteki çalışmaların bulanık sayılarla yapılması önerilmiştir. İlk kez Azerbaycan'lı bilim insanı Lütfü Askerzade tarafından bulanık mantık teorisi temelinde tanımlanmış bulanık kümeler kavramına göre ifade edilen bulanık sayılar belirsizliği ölçmede en uygun sayılar olarak ifade edilmektedir (Şen, 2001). Bu çalışmanın devamında yapılabilecek diğer bir çalışma ise farklı bilim alanlarını – Matematik, Fizik, Biyoloji, Coğrafya gibi- kapsayan araştırmalarla bu çalışmada önerilen yaklaşımın geçerlilik ve güvenilirlik boyutunun geliştirilmesidir.

References

- Akbaş, Y., & Gençtürk, E. (2013). Coğrafya öğretmenlerinin alternatif ölçme değerlendirme teknikleri ile ilgili görüşleri: Kullanma düzeyleri, sorunlar ve sınırlılıklar [Geography teachers' views about alternative assessment and evaluation techniques: usage levels problems and limitations]. *Eastern Geographical Review*, 18(30), 331-356.
- Altun, A., & Gelbal, S. (2014). Öğretmenlerinin kullandıkları ölçme ve değerlendirme yöntem veya araçlarının ikili karşılaştırma yöntemiyle belirlenmesi [Determining teachers' measurement tools or techniques via pairwise comparison method]. *Journal of Measurement and Evaluation in Education and Psychology*, 5(1), 1-11.
- Anıl, D., & Acar, M. (2008). Sınıf öğretmenlerinin ölçme değerlendirme sürecinde karşılaştıkları sorunlara ilişkin görüşleri [Elementary school teachers' views on issues they experience through measurement and evaluation processes]. *YYU Journal of Education Faculty*, 5(2), 44-61.
- Arı, A. (2010). Öğretmenlere göre proje ve performans görevlerinin uygulanmasında karşılaşılan sorunlar [Difficulties which teachers face in implementing projects and performance tasks]. *Electronic Journal of Social Sciences*, 9(34), 32-55.
- Baki, A., & Birgin, O. (2002). Matematik eğitiminde alternatif bir değerlendirme olarak bireysel gelişim dosyası uygulaması [Individual development file application as an alternative assessment in mathematics education]. *In proceedings of the 5th National Science and Mathematics Education Congress 16-18 September 2002*, Vol 2, (pp.913-920). METU, Ankara, Turkey.
- Bertiz, H., & Uluçınar Sağır, Ş. (2014). Fen ve teknoloji eğitiminde portfolyo ile zenginleştirilmiş öğrenme ortamı ve etkileri [The effects of learning environment enriched via portfolio in science and technology education]. *Journal of Turkish Science Education*, 11(2), 63-84.
- Birgin, O. (2008). Alternatif bir değerlendirme yöntemi olarak portfolyo değerlendirme uygulamasına ilişkin öğrenci görüşleri [Students' views about the application of portfolio assessment as an alternative assessment method]. *Journal of Turkish Science Education*, 6(1), 1-24.
- Büyüktokatlı, N., & Bayraktar, Ş. (2014). Fen eğitiminde alternatif ölçme değerlendirme uygulamaları [Alternative Assessment Practices in Science]. *Pegem Journal of Education & Instruction*, 4(1), 103-106.
- Chen, J-F., Hsieh, H-N., & Do, Q.H. (2015). Evaluating teaching performance based on fuzzy AHP and comprehensive evaluation approach. *Applied Soft Computing*, 28, 100-108.
- Çiftçi, S. (2010). İlköğretim birinci kademe 4. ve 5. Sınıf öğretmenlerinin performans görevlerine ilişkin görüşleri [The opinions of the teachers in upper primary classes concerning the student performance tasks]. *Elementary Education Online*, 9(3), 934-951.
- Dağdeviren, M., Yavuz, S., & Kılınç, N. (2009). Weapon Selection using the AHP and TOPSIS methods under Fuzzy Environment. *Expert Systems with Applications*, 36, 8143-8151.
- Dağhan, G., & Akkoyunlu, B. (2014). Bilişim teknolojileri dersinde kullanılan performans dayalı değerlendirme yöntemlerine ilişkin nitel bir çalışma [A qualitative study about performance based assessment methods used in information technologies lesson]. *Educational Sciences: Theory & Practice*, 14(1), 323-339.
- Demirbaş, M., & Yağbasan, R. (2004). Fen bilgisi öğretiminde duyuşsal özelliklerin değerlendirilmesinin işlevi ve öğretim süreci içinde öğretmen uygulamalarının analizi üzerine bir araştırma [A research on the progress of evaluating affective characteristics in science teaching and the analysis of teachers' practices in teaching process]. *Journal of Gazi University Kırşehir Education Faculty*, 5(2), 177-193.
- Deniz, E., & Kaptan, F. (2011). Yapılandırmacı fen öğretiminde tamamlayıcı ölçme değerlendirme uygulamalarından performans temelli değerlendirmenin önemi [The importance of performance-based assessment, in constructivist science education]. *Karadeniz Dergisi, Black Sea-Черное море*. 9, 25-43. Retrieved from <http://dergipark.gov.tr/kdeniz/issue/16866/175573>
- Döş, B. (2016). Öğretimi geliştirmede alternatif ölçme değerlendirme uygulamalarının analizi: Literatür taraması [Analysing The Alternative Assessment Applications for the Development of Teaching: Review of Literature]. *International Online Journal of Educational Sciences*, 8(4), 215-228.

- Dursun, M.F., Gülbay İ., Çetin S., & Tek Ü. (2012). *Orta öğretim 9. sınıf kimya ders kitabı (5.bs) [The 9th grade chemistry textbook of high school (5.Ed)]*. Ministry of National Education Books.
- Ersöz, F., Kabak, M., & Yılmaz, Z. (2011). Lisansüstü öğrenimde ders seçimine yönelik bir model önerisi [A model proposal for course selection for postgraduate students]. *Journal of the Faculty of Economics and Administrative Sciences*, 13(2), 227-249.
- Expert Choice. (2000). *Analytical Hierarchy Process (AHP) Software, Version 9.5, Expert Choice*, Pittsburgh.
- Gelbal, S., & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar [Teachers' proficiency perceptions of about the measurement and evaluation techniques and the problems they confront]. *H. U. Journal of Education*, 33, 135-145.
- Gupta, K. P., & Chaudhary, N. S. (2017). Prioritizing the Factors influencing Whistle Blowing Intentions of Teachers in Higher Education Institutes in India. *Procedia Computer Science*, 122, 25-32.
- Güler, N. (2017). *Eğitimde ölçme ve değerlendirme (10. bs.) [Measurement and evaluation in education (10 th ed.)]* Ankara: Pegem Academy Publishing.
- Güntut, M., Güneş, P., & Çetin, S. (2017). *Orta öğretim 9. Sınıf kimya ders kitabı (1.bs) [The 9th grade chemistry textbook of high school (1. Ed)]*. Ministry of National Education Books.
- Hwang, C. L., Yoon, K. (1981). *Multiple attribute decision making: Methods and applications, a state of the art survey*. New York: Springer-Verlag.
- Kaya, S., Karaçam, S., Eş, H., & Tuncel, M. (2013). 4. ve 5. Sınıf öğretmenlerinin fen ve teknoloji dersinde proje ve performans görevlerine ilişkin görüşleri [The 4th and 5th grade primary teachers' opinions related to project and performance tasks at science and technology course]. *Pamukkale University Journal of Education*, 33, 187-201.
- Kuran, K., & Kanatlı, F. (2009). Alternatif ölçme değerlendirme teknikleri konusunda sınıf öğretmenlerinin görüşlerinin değerlendirilmesi [The evaluation of classroom teachers' opinions on the alternative assessments techniques]. *Mustafa Kemal University Journal of Social Sciences Institute*, 6(12), 209-234.
- Lai, J. H., & Choi, E. C. (2015). Performance measurement for teaching hotels: a hierarchical system incorporating facilities management. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 16, 48-58.
- Lupo, T., (2013). A fuzzy ServQual based method for reliable measurements of education quality in Italian higher education area. *Expert Systems with Applications*, 40(17), 7096-7110.
- MEB. (2013). *Millî Eğitim Bakanlığı Orta Öğretim Kurumları Yönetmeliği. Resmî Gazete [Ministry of National Education Regulation on Secondary Education Institutions. Official newspaper]*. Number: 28758, 07.09.2013. Retrieved 16 December 2016, from https://ogm.meb.gov.tr/meb_iys_dosyalar/2016_11/03111224_ooky.pdf.
- MEB. (2007). *Ortaöğretim 9. Sınıf Kimya Dersi Öğretim Programı [The 9th grade chemistry curriculum of high school]*. Ankara: Presidency of Education and Training Board.
- MEB. (2017). *Ortaöğretim Kimya Dersi Öğretim Programı [Secondary School Chemistry Curriculum]*. Ankara: Ministry of National Education.
- Metin, M. (2013). Öğretmenlerin performans görevlerini hazırlarken ve uygularken karşılaştığı sorunlar [Teachers' difficulties in preparation and implementation of performance task]. *Educational Sciences: Theory & Practice*, 3(3), 1645-1673.
- Metin, M., & Özmen, H. (2010). Fen ve teknoloji öğretmenlerinin performans değerlendirmeye yönelik hizmet içi eğitim (HİE) ihtiyaçlarının belirlenmesi [Determination of science and technology teachers' in-service education (INSET) needs on performance assessment]. *Kastamonu Education Journal*, 18(3), 819-838.
- Nazlıççek, N., & Akarsu, F. (2008). Fizik, kimya ve matematik öğretmenlerinin değerlendirme araçlarıyla ilgili yaklaşımları ve uygulamaları [Physics, chemistry and mathematics teachers' approaches to assessment tools and their assessment practices]. *Education and Science*, 33(149), 18-29.

- Özçelik, D.A. (1992). *Ölçme ve değerlendirme [Measurement and evaluation]*. Student Selection and the Placement Centre (ÖSYM) Publications, Ankara.
- Özdemir, S.E. (2010). İlköğretim öğretmenlerinin alternatif ölçme ve değerlendirme araçlarına ilişkin yeterlikleri ve hizmet içi eğitim ihtiyaçları [Elementary teacher competencies and in service training needs in alternative measurement and assessment tools]. *The Journal of Turkish Educational Sciences*, 8(4), 787-816.
- Saaty, T.L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill International Book Company. USA.
- Saaty, T.L., (1986, Axiomatic Foundation of the Analytic Hierarchy Process. *Management Science*, 32(7), 841-855.
- Saaty, T.L., (1991, Some Mathematical Concepts of the Analytic Hierarchy Process. *Behaviormetrika*. 29, 1-9.
- Sağlam-Arslan, A., Devocioğlu-Kaymakçı, & Y., Arslan, S. (2009). Alternatif ölçme-değerlendirme etkinliklerinde karşılaşılan problemler: fen ve teknoloji öğretmenleri örneği [Problems concerning alternative evaluation methods: The case of science and technology teachers]. *Ondokuz Mayıs University Journal of Education Faculty*, 28, 1-12.
- Shyur, H-J. (2006). COTS evaluation using modified TOPSIS and ANP. *Applied Mathematics and Computation*, 177, 251-259.
- Şenel Çoruhlu, T., Er Nas, S., & Çepni, S. (2009). Fen ve teknoloji öğretmenlerinin alternatif ölçme-değerlendirme tekniklerini kullanmada karşılaştıkları problemler: Trabzon örneği [Problems facing science and technology teachers using alternative assessment technics: Trabzon sample]. *YYU Journal of Education Faculty*, 6(1), 122-141.
- Şen, Z. (2001). *Bulanık mantık ve modelleme ilkeleri [Fuzzy Logic and Modeling principles]*. İstanbul: Bilge Kültür Sanat.
- T.C. Resmî Gazete. (2017). *Millî Eğitim Bakanlığı Ortaöğretim Kurumları Yönetmeliğinde Değişiklik Yapılmasına Dair Yönetmelik. [Regulation on the Amendment of the Regulation of the Secondary Education Institutions of the Ministry of National Education]*. No: 30182, 16.09.2017. Retrieved 16 December 2016, from <http://www.resmigazete.gov.tr/eskiler/2017/09/20170916.pdf>.
- Wu, H-Y., Chen, J-K., Chen, I-S., & Zhuo, H-H. (2012). Ranking universities based on performance evaluation by a hybrid MCDM model. *Measurement*, 45, 856-880.
- Yaşar, M.D., & Sözbilir, M. (2013). Öğretmenlerin 2007 yılı kimya dersi öğretim programındaki yapılandırmacılığa dayalı öğelere yönelik algılamaları [Teachers' perceptions about constructivist principles in the 2007 chemistry curriculum]. *Journal of Turkish Science Education*, 10(4), 75-102.
- Yıldız, İ., & Uyanık, N. (2004). Matematik eğitiminde ölçme-değerlendirme üzerine [On measurement and evaluation in mathematics teaching]. *Gazi University Kastamonu Education Journal*, 12(1), 97-104.
- Yiğit, F., & Kırımlı, B. (2015). Türkçe öğretmenlerinin alternatif ölçme değerlendirme yöntemlerinin işlevleri ve kullanma sıklığı hakkındaki görüşleri [Teacher opinions on functions and frequency of using of alternative assessment and evaluation methods]. *National Education*, 44(205), 64-85.
- Yüksel, M. (2011). Eğitim ve öğretim kazanımları temelinde 9. Sınıf kimya ders kitabının incelenmesi. [Investigation of the 9th grade chemistry course book on the basis of the education and teaching acquisitions]. *Selçuk University Journal of Ahmet Keleşoğlu Education Faculty*, 32, 29-48.
- Yüksel, İ., & Dağdeviren, M. (2006). Sosyo-teknik sistemlerde hatalı davranış riskini belirlemeye yönelik bir erken uyarı modeli. [An early warning model to identify faulty behaviors risk in social-technique systems and its application] *Journal of The Faculty of Engineering and Architecture of Gazi University*, 21(4), 791-799.
- Yüksel, M., & Geban, Ö. (2015). Kimya öğretmeni özel alan yeterliklerine göre öğretmen performansının değerlendirilmesi [Evaluation of teacher performance according to the special area competencies of chemistry teachers]. *H. U. Journal of Education*, 30 (1), 299-312.
- Zyoud, S.H., & Fuchs-Hanusch, D. (2017). A bibliometric-based survey on AHP and TOPSIS techniques. *Expert Systems with Application*, 78, 158-181.