



## Examining Turkish Course Exam Questions in Terms of Originality, Page Layout, Item Type, Item Writing Criteria and Cognitive Level \*

Pınar Kanık Uysal <sup>1</sup>, Çiğdem Akın Arıkan <sup>2</sup>, Tuba Acar Erdol <sup>3</sup>,  
Emel Bayrak Özmutlu <sup>4</sup>, Hayati Akyol <sup>5</sup>

### Abstract

The aim of this study was to examine Turkish course exam questions in terms of originality, page layout, item types, item writing criteria and cognitive level. The research was conducted with the document analysis method, and to determine the data source, the maximum variation sampling method, one of the purposive sampling methods, was used. The data source of the research consisted of exam questions prepared by Turkish teachers employed in secondary schools in a medium-sized city in the Black Sea region. Exam questions from each of the school levels determined on the basis of High School Entrance Exam achievements, from each district within the provincial borders, and from each type of school at the secondary school level were included in the research. Within the scope of the research, 747 questions in 51 exam papers were analysed. Descriptive analysis was carried out for the research data, and the obtained findings were summarised using frequencies and percentages. When the research findings were examined, it was concluded that the exam papers had similarity rates ranging from 11% to 91%, and that several exam papers and questions were taken from different internet sources as they were without making any changes. Other obtained findings were that more than half of the Turkish course exam papers did not have the page layout features that should be found in an exam paper, that the most frequently used item type in the exam papers was the multiple-choice item type, and that teachers did not pay attention to the item writing criteria. Another striking finding of the study was that about four-fifths of the questions examined were prepared for the *understand* level and one-fifth for the *remember* and *apply* level, and that while very few questions were prepared for the *analyze* and *create* levels, no questions were prepared for the *evaluate* level. Moreover, it was

### Keywords

Turkish course  
Measurement and evaluation  
Similarity  
Item types  
Revised Bloom Taxonomy

### Article Info

Received: 05.11.2021  
Accepted: 04.01.2022  
Online Published: 04.29.2022

DOI: 10.15390/EB.2022.10896

\* This article was supported by Ordu University's Scientific Research Commity under the code A-2027. Additionally, it is orally presented at the 8th International Congress on Curriculum and Instruction.

<sup>1</sup> Ordu University, Faculty of Education, Dept. of Turkish and Social Sciences Education, Turkey, [pinaruysal32@gmail.com](mailto:pinaruysal32@gmail.com)

<sup>2</sup> Ordu University, Faculty of Education, Dept. of Educational Sciences, Turkey, [akincgdm@gmail.com](mailto:akincgdm@gmail.com)

<sup>3</sup> Ordu University, Faculty of Education, Dept. of Educational Sciences, Turkey, [tubaacarerdol@gmail.com](mailto:tubaacarerdol@gmail.com)

<sup>4</sup> Ordu University, Faculty of Education, Dept. of Primary Education, Turkey, [emelbayrakozmutlu@gmail.com](mailto:emelbayrakozmutlu@gmail.com)

<sup>5</sup> Gazi University, Faculty of Education, Dept. of Primary Education, Turkey, [hayatiakyol@gmail.com](mailto:hayatiakyol@gmail.com)

concluded that the majority of the questions examined were prepared for the middle cognitive level, while higher-order questions were included only at a rate of eight per thousand. The obtained results are discussed within the scope of the literature and a number of recommendations have been made.

## Introduction

Learning and teaching is a process that requires constant information gathering, decision making (Nitko & Brookhart, 2016) and planning the next step according to the feedback received. Measurement and evaluation, which is an integral part of this process, is used to determine the extent to which the knowledge and skills intended to be fostered by the curricula are acquired (Ministry of National Education [MoNE], 2019), and the measurement and evaluation studies that are conducted form the basis of the decisions regarding the plans to be made and the steps to be taken. It is accepted that evaluation, which is a method of systematically gathering, interpreting and using information, improves learning and that a good evaluation is an indicator of good teaching (Döş et al., 2016; McMillan, 2015; Tofade, Elsner, & Haines, 2013). Evaluation, which is an integrated part of learning and teaching, is carried out before, during and after instruction, and teachers make evaluations for the purposes of identification, shaping learning and level determination. Evaluations made before instruction reveal students' prior knowledge and guide the planning of in-class activities; situations such as structured exercises carried out in the classroom, interim assessments, in-class tasks, and observations offer the opportunity to shape the learning process; mid-term and end-of-term exams provide the opportunity make evaluations aimed at level determination. The conducted studies (Belet & Sağlam, 2015; Beyhan, 2012; Deniz & Keray Dinçel, 2019; Göçer, 2005; Güneylü & Abbasoğlu, 2015; Karadüz, 2009) show that among the types of evaluation, the assessment type most frequently preferred by teachers is level determination. In this type of assessment, test-type measurement tools are generally used, and since these are more objective and standardised than other evaluation techniques (observation, checklists, rubrics, etc.), they prevent some of the subjectivity and inconsistency that affect evaluation (Nitko & Brookhart, 2016). Teachers use this type of assessment to determine the extent to which students have achieved the learning goals they need to achieve following instruction, and in addition, to decide which thinking skills students are able to exhibit at the end of the process.

Deciding which thinking skills teachers want their students to use to answer questions also requires teachers to use a classification while preparing the questions. For this reason, taxonomies are used while preparing questions, and taxonomies enable the use of a common language by offering a systematic approach for defining the nature of what is learned (McMillan, 2015). The use of taxonomy draws attention to the diversity of abilities and skills that can be managed in instruction and assessment (Nitko & Brookhart, 2016), and this also directs the curricula and teachers in determining and evaluating learning goals in such a way that students can reach each of the levels in the taxonomies. In the literature, there are many taxonomies (Barrett, Haladayna, Marzano, etc.) for the graded classification of learning goals. Among these, Bloom's Taxonomy stands out as the best known and most widely used taxonomy (Büyükalın-Filiz & Yıldırım, 2019; Lipscomb, 1985; McMillan, 2015; Nitko & Brookhart, 2016; Tofade et al., 2013). This taxonomy, which was revised by Anderson and his team in 2001 (Anderson et al., 2001), enables educators to develop a more complete understanding of specific goals and allows them to use this understanding to develop the basic link between assessment and instruction (Anderson, 2005). By means of this taxonomy, various learning objectives, and exam questions for these objectives, can be prepared. Questions prepared on the basis of a classification not only provide convenience to the teacher for teaching, but also improve students' higher-order thinking skills (Uymaz & Çalışkan, 2019) and prevent questions from accumulating in certain levels (Büyükalın, 2007). High-order thinking skills are affected by the level of the questions used in educational environments (Aslan, 2011). The use of a taxonomy while preparing exam questions provides a tool for deciding whether higher-order or low-order thinking skills are taught and evaluated (Nitko & Brookhart, 2016). In order to carry out a good

quality measurement and evaluation process, teachers should possess not only cognitive taxonomy knowledge but also adequate knowledge about validity and reliability, which are two basic features that should be included in measurement tools. Adequate reliability is required for valid measurement results, and reliability can be ensured for exams in which a sufficient number of questions of appropriate difficulty are included, clear and understandable instructions and question items are found, and scoring rubrics are prepared, and this also indicates the consistency of the evaluation results (Nitko & Brookhart, 2016). Lack of time (insufficient time given in the exam), and the fact that the exam questions are too difficult or too easy, that the questions do not include different levels of cognitive taxonomy, and that they do not cover the topics or outcomes taught in the course have a negative effect on the validity of the exam. Considering all these aspects, teachers need to have the competence to prepare and implement measurement and evaluation tools that are appropriate for their field and for the developmental characteristics of their students in order to be able to make a valid and reliable assessment in accordance with the learning objectives and thinking skills at their own grade level (MoNE, 2017).

Competence in the field of measurement and evaluation, which is one of the most important professional competencies of teachers (Esen, 2019; MoNE, 2006, 2019; McMillan, 2015; Nitko & Brookhart, 2016; Xu & Brown, 2016) gives teachers the responsibility to decide what is to be measured and to determine the appropriate question type according to the purpose (Kubiszyn & Borich, 2003). However, studies conducted on teacher competence (Benzer & Eldem, 2013; Deniz & Keray Dınçel, 2019; Esen, 2019; İnceçam, Demir, & Demir, 2018; Gelbal & Kelecioğlu, 2007; MoNE, 2006; Yaşar, Gültekin, Türkkkan, Yıldız, & Girmen, 2005) reveal that the area in which teachers regard themselves as least competent is that of measurement and evaluation. When the studies made in this context are examined, Mertler (2004), in his study aimed at determining the assessment literacy of secondary school teachers and teacher candidates, determined that the area in which both the teachers and teacher candidates were most inefficient was that of developing a valid measurement tool. In the study they carried out to determine primary and secondary school teachers' perceptions of measurement and evaluation efficacy, Gelbal and Kelecioğlu (2007) concluded that teachers needed in-service training in the preparation and use of measurement techniques, while Karacaoğlu (2008) concluded that teachers found themselves less competent in using different measurement methods, and Esen (2019) concluded that the area in which teachers considered themselves the least competent was "writing open-ended items that can measure higher-order intellectual skills". The study conducted by İnceçam et al. (2018) with secondary school teachers revealed that teachers were incompetent in the whole process of writing open-ended items, and as reasons for this, they listed reasons such as their own lack of knowledge, low student level, inappropriate subjects for writing open-ended items, and the long time taken for open-ended item writing and scoring. Özen (2020) determined that Turkish teachers had problems in preparing and scoring open-ended questions, and that as reasons for this, they cited lack of time, and problems experienced in content validity and evaluation. It is a professional responsibility for teachers to have knowledge of the measurement and evaluation field and to reflect this in practice, while it is a situation that is contrary to professional standards, unethical and even illegal for them to make poor quality or erroneous assessments and to harm their students with the incorrect decisions they make as a result of these assessments (Nitko & Brookhart, 2016).

Another responsibility of teachers who have responsibility for creating their own assessment methods is to choose assessment methods that have been prepared by others but serve the purposes of the lesson. Teachers are obliged to check the questions that they did not prepare themselves and that they obtained ready-made as to whether they are suitable for their own teaching purposes, and to take all precautions so as not to make an erroneous assessment. Just as teachers have responsibilities to ensure the validity and reliability of the questions they prepare themselves, the same is true for the questions they obtain ready-made. Nitko and Brookhart (2016) state that teachers who do not subject ready-made questions to their own criteria have a high probability of making an erroneous assessment, that ready-made questions may not match the content emphasised by the teacher in his/her own lessons, and that therefore, the questions that are obtained should be examined and checked by the teachers

themselves in terms of appropriateness for the course content and assessment criteria. For this reason, rather than using ready-made questions as they are, teachers should adapt them in accordance with the measurement and evaluation principles, use the ones that are appropriate for the outcomes they plan to measure, and choose the question type that is most suitable for the purpose by knowing the strengths and weaknesses of the question types. The types of questions to be selected vary according to the purpose, area and thinking skills to be measured. Question types vary as true-false, matching, short-answer, fill-in-the-blank, multiple-choice, open-ended, etc. (Baykul, 2010; Hogan, 2007; Popham, 2005). Making use of a variety of measurement tools according to the outcome or characteristic intended to be measured instead of using a single question type also enables a more valid measurement to be made (Deniz & Keray Dinçel, 2019).

While structuring exams, making use of various item types that allow students to use their higher-order cognitive skills (Aslan, 2011; MoNE, 2019; Tofade et al., 2013) necessitates the preparation of questions that require students to experience a production and creation process in a real sense rather than simply answering questions, by choosing learning objectives at complex thinking levels. According to Bloom's Taxonomy, it is emphasised that the learning objectives in the first three cognitive categories (remember, understand and apply) can be evaluated with short-answer, true-false, multiple-choice or matching test techniques, while the learning objectives in the last three cognitive categories (analyze, evaluate and create) can only be partially measured with test techniques such as these, and that therefore, evaluations should be made with performance-based assessment methods (projects, performance, portfolios) (Nitko & Brookhart, 2016), since it is accepted that knowledge and basic comprehension goals can be better evaluated with multiple-choice and short-answer items, while in-depth understanding and reasoning goals can be better assessed with open-ended items (McMillan, 2015). The Turkish course also has the characteristic of a course that is suitable for preparing open-ended questions, the inclusion of different thinking skills, and making use of various question types. Scarino (2013) argues that language teachers are the most important of all stakeholders, since they are direct test users. This situation places great responsibility on Turkish teachers in the area of measurement and evaluation competence, and they are expected to prepare valid and reliable questions at different cognitive levels. Considering that success in the Turkish course affects the achievement level in all other courses (Benzer, 2019; Ceran & Deniz, 2015; Karatay & Dilekçi, 2019; Yıldız et al., 2019), the importance of the need to conduct studies examining the appropriateness of the questions used in this course in terms of measurement and evaluation criteria, and determining and improving the quality of the questions applied in the lessons, becomes evident.

When the literature is examined, it is seen that studies have been carried out on teachers' knowledge and competencies in measurement and evaluation (Benzer & Eldem, 2013; İnceçam et al., 2018; Özen, 2020; Sultana, 2019; Xu & Brown, 2016; Wilen, 1991), questioning strategies (Anisah, Fitriati, & Rukmini, 2019; Suartini, Wedhanti, & Suprianti, 2020) and question levels (Döş et al., 2016; Tofade et al., 2013). In the studies conducted specifically for the Turkish course and the exams made for this course, it is seen that they have generally been conducted on question levels (Akyol, Yıldırım, Ateş, & Çetinkaya, 2013; Bekaroğlu, 2007; Çintaş-Yıldız, 2015; Göçer, 2005, 2016; Güfta & Zorbaz, 2008; Güneyleli & Abbasoğlu, 2015; Kavruk & Çeçen, 2013; Ülger, 2003), question types (Aydın & Uçgun, 2020; Maden & Durukan, 2009), question areas (Aydın & Uçgun, 2020; Güfta & Zorbaz, 2008; Maden & Durukan, 2009), competence in question preparation (Karatay & Dilekçi, 2019; Özen, 2020), and the extent to which the questions are appropriate for the outcomes in the curriculum (Aydın & Uçgun, 2020; Maden & Durukan, 2009). However, no study was found examining whether or not the Turkish course exam questions were taken directly from internet sources, the page layouts of the exam papers, the appropriateness of each question type for its specific item writing rules and the distribution of the item types together with these, and the distribution of items according to the cognitive process dimension and metacognitive level of the Revised Bloom Taxonomy (RBT). In this respect, the aim of this study is to reveal how adequate Turkish exam papers are in terms of including original and different types of questions and questions that require students to use their higher-order thinking skills. For this purpose, in the study, an attempt is made to present a holistic evaluation by including in the sample the

examination papers from each of the school levels determined on the basis of High School Entrance Exam (LGS) achievements, from each district within the provincial borders, and from each type of school at the secondary school level. In this way, the aim is to obtain a comprehensive and deep description of the quality and adequacies of Turkish course exams in terms of measurement and evaluation.

For this purpose, answers were sought to the following questions:

1. What are the similarity rates of Turkish course exam papers in terms of content?
2. What is the distribution of Turkish course exam papers in terms of page layout?
3. What is the distribution of question types in Turkish course exam papers?
4. Are the necessary rules observed while writing the questions in Turkish course exam papers?
5. What is the distribution of the questions in Turkish course exam papers according to the cognitive process dimension and metacognitive level of the RBT?

## **Method**

### ***Research Model***

In this study examining Turkish course exam questions, the document analysis method was used. Document analysis is the systematic examination of existing resources to access verifiable data (Watkins, Meiers, & Visser, 2012). This means examining documents related to the subject under investigation in accordance with scientific principles. In this method, a systematic process is followed that requires the examination and evaluation of sources that are considered as documents (e.g., books, diaries, newspapers, journals, exam papers, etc.). As with other qualitative research methods, this method also requires the examination and interpretation of data in order to obtain an understanding about the subject studied (Corbin & Strauss, 2008). Since the aim of this study is to obtain a description of the qualities and adequacies of Turkish course exams in terms of measurement and evaluation, it was decided that it would be appropriate to adopt the document analysis method in the research.

### ***Data Source***

The data source of this study consists of exam papers prepared by Turkish teachers working in secondary schools in a medium-sized province in the Black Sea region. While determining the sample of the study, the maximum variation sampling method, which is one of the purposive sampling methods, was used. The purpose of maximum variation sampling is not to generalise, but, by sampling different situations related to the problem, to provide the researcher with richer information about the situations (Büyüköztürk, Kılıç Çakmak, Akgün, Karadeniz, & Demirel, 2012). In order to ensure maximum variation within the scope of the research, exam papers from different school levels (low, medium and high achievement levels), different districts (17 districts in the province where the data were collected), and different school types (secondary schools, *imam hatip* secondary schools) were included in the sample. There are a total of 199 secondary schools in the specified province. Three of these schools were not included in the study because they were newly opened and there were no students taking the High School Entrance Exam (LGS), while 10 of them were not included because they are private schools. Accordingly, in line with the information received from the Provincial Directorate of National Education, the secondary schools included in the research were divided into three levels as low, medium and high according to their success in the LGS Turkish course in the 2018-2019 academic year. Then, taking these levels into account, schools from each level were classified according to the districts and school types they belonged to, and a total of 54 schools were selected. As a result of the report of the Turnitin program, since the exam papers of four of these schools were exactly the same, the written papers of three schools were excluded from the analysis, and the analysis was performed on the exam papers of 51 schools. The first and second exam papers prepared by eighth grade Turkish teachers employed in these schools were collected, but only the first exam papers were included in the analysis. In the selection of the first exam papers, having two exams in the Turkish course in one semester and the use of different question types in the first exam were effective. It was observed that in their first exams, the Turkish teachers used open-ended, fill-in-the-blank, matching, multiple-choice,

etc., item types, but that they mostly preferred multiple-choice item types in their second exams. For this reason, in accordance with the purpose of the research, the first exams containing different item types were included in the analysis.

### *Data Collection*

Ethics committee approval for this study was obtained from the Social and Human Sciences Research Ethics Committee of Ordu University (dated 03.11.2020 and numbered 2020-26). The data source of the research consists of “exam questions and answer keys” prepared by Turkish teachers for the eighth grade level Turkish course in the first semester of the 2019-2020 academic year. After the necessary permission was received from the Provincial Directorate of National Education, the exam papers were obtained through the Measurement and Evaluation Centre (MEC). In the first stage, the incoming data were examined by the researchers by preparing a checklist, and missing documents were again requested from the schools via the MEC and delivered to the researchers.

### *Data Analysis*

In the first sub-problem of the research, it was aimed to determine the similarity ratios of the Turkish course exam papers in terms of content. Within this scope, the Turnitin program was used. The Turnitin program is considered to be a suitable tool for measuring the originality of the exam papers, as it is a program that shows the non-original parts of the exam papers, the percentage of these parts, and the source from which the relevant section is obtained. In the analysis phase, each exam paper was uploaded to the Turnitin program in PDF file format with the name of the school. A report was created by the program showing the similarity rate, content and sources for each exam paper uploaded to the system. In the analysis of the reports, a four-category system was used to examine the exam papers on the basis of the similarity ratios. In this system, papers with a 0-24% similarity rate were classified as fourth level, 25-49% as third level, 50-74% as second level, and 75-100% as first level. Then, information about each category was shown by the descriptive statistics of frequencies and percentages. In addition, based on the Turnitin reports, it was determined how many different internet addresses were used in total by counting each similar internet address and the similarity rates, and the ones with the highest usage rate were determined.

The aim of the second sub-problem of the research was to examine the Turkish course exam papers in terms of form, and in this direction, the page layout was examined. For the page layout, descriptive analysis was used by considering criteria such as font type, letter size, adequate spacing between items and options, and item roots and options being on the same page, and evaluation was made using frequencies and percentages.

In the third and fourth sub-problems of the research, it was aimed to examine the distribution of the item types in the Turkish course exam papers and whether the necessary rules were observed while writing the items, respectively. When the exam papers were examined, it was seen that multiple-choice, true-false, open-ended, short-answer and matching test items were used. By examining the literature, the criteria that item writers/teachers should comply with in writing items for each item type were written (Atılğan, 2006; Baykul, 2010; Doğan, 2019a, 2019b; Haladyna, 1994, 1997; Miller, Linn, & Gronlund, 2009; Özçelik, 1992; Tekin, 2012; Thorndike, 1971). The specified criteria were sent to two measurement and evaluation experts, and accordingly, the necessary analyses were made by giving the criteria their final form. The criteria to be observed for each item type in the Turkish course exam papers were evaluated as “yes” and “no”. The teachers’ question-writing competencies were examined by the researchers and the results were interpreted as frequencies and percentages.

The thematic content analysis technique was used for examination of the questions in the exam papers, which constitute the fifth sub-problem of the research. As themes, the levels in the cognitive process dimension of the 2001 Revised Bloom Taxonomy (RBT) were evaluated as the unit of analysis. A total of 748 questions in 51 exam papers were examined, and since there was no question root in one of these questions, 747 questions were analysed. In the analysis of the questions in the written exams according to the RBT, the following steps were followed and the criteria were taken into account:

- Sample questions in studies conducted according to the RBT in the field of Turkish (Aktaş, 2017; Çintaş-Yıldız, 2015; Erdoğan, 2017; Göçer, 2016; Güneyli & Abbasoğlu, 2015; Kavruk & Çeçen, 2013; Sallabaş & Yılmaz, 2020) were examined. In these studies, the distribution of the given outcomes or sample questions in the taxonomy was evaluated by the researchers, and it was concluded that it was not sufficient to consider the question item only in terms of the outcome or subject area when deciding on the cognitive level, and that the type of question used and the task expected from the student were effective on the cognitive level of the question. For this reason, it was determined that there was a need for expert opinion regarding the level(s) at which the questions prepared for the same subject area and outcome but in different question types, would be coded.
- Videoconferences were held on three occasions with the participation of three experts and four researchers with higher-order and context-based experience in question writing in the field of Turkish education, and criteria were created to determine under which cognitive level the questions in different subject areas (e.g., main idea, narration styles, elements of the sentence, ways of developing the idea, punctuation/spelling rules, subjective/objective judgments, and rhetoric) would be coded. For example, questions about the main idea were coded at the *understand* level [see Table 1, Question 5]. The reason for this is that the main idea in the examined questions remained within the boundaries of the text and that in some texts, it was emphasised by writing it in italics, bold or quotation marks at the end of the text. Questions about narration styles were coded at the *remember* and *understand* level [see Table 1, Question 2 and Question 6]. The reason for this is that, for example, the second question, coded at the *remember* level, is at a level that can only be solved by recalling existing knowledge. The reason why the fifth question was coded at the *understand* level is that it requires the transfer of pre-existing knowledge in line with the given text. In questions related to the elements of the sentence, the complex ones (due to requiring a higher cognitive load) were evaluated at the *analyze* level [see: Table 1, Question 10], while the others were evaluated at the *apply* level. The reason for evaluation at the *apply* level is that the analysis of the sentence requires a certain sequence of operations, and after the predicate is determined, a number of questions are asked to find the subject, object and complement to the predicate, respectively. The reason for evaluation at the *analyze* level is that it is difficult to determine the structure in complex sentences, and finding the elements of the sentence requires dividing this structure into meaningful parts. Questions on ways of developing the idea were coded at the *understand* level [see Table 1, Question 7]. The reason for this is that the ways of developing the idea are divided into types such as defining, exemplifying, utilising numerical data, and giving evidence. In the questions that were examined, a text was generally given and it was asked which way(s) of developing the idea were/were not included in this text. These questions require students to find a solution by making connections between their existing knowledge about ways of developing the idea and the information given in the text. Questions about punctuation marks/spelling rules were coded at the *remember*, *understand* and *apply* levels [see Table 1, Question 3, Question 8 and Question 9]. The reason for encoding at the *remember* level is the necessity for the student to recall information directly from memory; the reason for encoding at the *understand* level is that the student is required to place what he/she knows about punctuation at appropriate points in the text, and the reason for encoding at the *apply* level can be shown as the fact that the student puts what he/she knows about punctuation into practice by using his/her writing skills.
- The analyses were conducted by three researchers in line with the criteria for Turkish questions and the table consisting of cognitive process groups, alternative names, definitions, relationships and differences prepared for the RBT by Acar-Erdol (2020).
- In cases where the analysis of the questions was undecided, the teachers' answer keys were examined and coding was done in line with the expected answer.

- The analyses were made in line with the cognitive process sublevels of the RBT (for example, the coding of the question was made not for the *apply* level, but for the *executing* and *implementing* steps, which form the sublevels of this level), while the presentation of the analysis results through frequencies and percentages was in line with the cognitive process levels (for example, *apply*). Sample questions coded in the cognitive process levels of the RBT as a result of the analyses are given in Table 1.

**Table 1.** Sample Questions Coded in the Cognitive Process Levels of the RBT

Cognitive Process Levels	Sample Questions							
Remember	1. Evaluate the following information* as true (T) or false (F). (2 marks) [ ] Verbal suffixes are added to noun roots and stems [ ] An idiom is a stereotyped phrase that has a specific meaning that is more or less different from its real meaning.							
	2. Match the following gaps with the appropriate expression (5 marks)  <table style="width: 100%; border: none;"> <tr> <td style="width: 60%;">1 It is written to change the reader’s mind</td> <td style="width: 40%;">Narration</td> </tr> <tr> <td>2 It is a way of depicting entities with words.</td> <td>Argumentation</td> </tr> <tr> <td>3 Relating the event in the past tense is dominant.</td> <td>Description</td> </tr> <tr> <td>4 The aim is to provide information on a subject.</td> <td>Explanation</td> </tr> </table> <p>3. A ..... is used to indicate uncertain or suspect information, unknown dates and places.**                      When writing dates, a ..... is used to separate numbers that represent the day, month, and year.                      A ..... is used to separate sequential sentences with a comma between them.</p>	1 It is written to change the reader’s mind	Narration	2 It is a way of depicting entities with words.	Argumentation	3 Relating the event in the past tense is dominant.	Description	4 The aim is to provide information on a subject.
1 It is written to change the reader’s mind	Narration							
2 It is a way of depicting entities with words.	Argumentation							
3 Relating the event in the past tense is dominant.	Description							
4 The aim is to provide information on a subject.	Explanation							
Understand	4. Separate the following sentences into purpose, reason, condition, and comparison (5 marks) [.....] We must be organised in order to make good use of our time. [.....] Increasing responsibilities require careful use of our time. [.....] It seems as though we have less time than in childhood. [.....] If a person is organised, he/she will use his/her time better. [.....] He went to the committee room in order to introduce himself.							
	5. When we were children, the days were endless, and we lived the longest times of our lives. We thought that we did everything and we did whatever we wanted in our spare time. (I) It is also a fact that time flows fast whether or not we want it to, because our responsibilities have increased. We have responsibilities towards our family, our job, our neighbour... (II) Since it is not possible to stop the flow of time, we have to appreciate every moment that we live. (III) While working for this, we should work for a planned and orderly life. We should do all our work on time, and we should not delay any of our work. Considering that many tasks, which are postponed with the thought that they will be done in the future, cannot be done and are neglected, the importance of the proverb, “Never put off till tomorrow what you can do today” is revealed once again. Write the message (main idea) that is intended to be given in the text (5 marks)							
	6. With the first light of morning, we set out with our friends to see the Ishak Pasha Palace. On the way there, some friends said that the palace and its surroundings							



were protected and guarded. When we arrived at the palace, we were disappointed. The walls around the palace were destroyed and had been breached in places. Its inscriptions were too faded to be read. The interior was covered with food scraps thrown by the shepherds and plastic bags.

Which of the following is used in the explanation of this passage? (4 marks)

- A) Description – Comparison B) Narration – Description  
C) Discussion – Explanation D) Comparison – Narration

Many famous names in the world have given advice on tolerance and struggled with this issue. Among these, Mevlana and Yunus Emre will never be forgotten.

7. Which of the following ways of developing the idea is used in the above text?

- A) Comparison B) Giving evidence C) Definition D) Exemplification

8. There are three things in life that are difficult ( ) keeping a secret ( ) forgetting an injury ( ) using one's free time ( )

Which punctuation marks should be placed in the blank parentheses above, respectively? (5 marks)

- A) (:) (,) (,) (.) B) (:) (;) (,) (... ) C) (;) (,) (,) (.) D) (.) (,) (,) (.)

Apply\*\*\* 9. Write an explanatory paragraph for the saying given below. Do not forget to give examples while giving your explanation. (15 marks)  
Relevance to the topic and main idea: 5 marks  
Spelling and punctuation: 5 marks  
Page layout: 5 marks

“Intelligence, like a field, needs to be planted and tended.” Cicero.

Analyze 10. Find the elements of the sentence below and write them under it. (5 marks)  
-The things people do during their lifetime are only half of the things they want to do.

Evaluate \*\*\*\*

Create 11. “The lazy person who does not take one step in time has to take a hundred steps later.....” Giovio  
Write a story using the words in brackets. (people, time, youth, work, past, lazy) (30 marks)

\*The spelling mistakes in the exam papers are given as they are, and no corrections have been made.

\*\*Although the root of the question is not included, it is understood from the given items that punctuation marks are asked for.

\*\*\*The reason why this question is given as an example is that punctuation marks and spelling rules are put into practice by using them in writing skills.

\*\*\*\* Since no question could be accessed in the evaluate level as a result of the analyses, this level could not be sampled.

In the evaluation of the exam papers according to the metacognitive level, by using the classification made by Moodley (2013), the “remember” level of the RBT was evaluated as the lower cognitive level, the “understand” and “apply” levels were evaluated as the middle cognitive level, and the “analyse”, “evaluate” and “create” levels were evaluated as the higher-order cognitive level, and the analyses were made accordingly. The distribution of the questions according to the cognitive level as a result of the analyses is presented through frequencies and percentages.

### ***Validity and Reliability***

In addition to the validity and reliability studies described under the headings in the Method section, three researchers (specialists in the Turkish subject area, measurement and evaluation, and curriculum development) worked together in order to strengthen the reliability of the study, and the results were re-evaluated by a fourth researcher (a curriculum development specialist). Three

researchers (a specialist in the Turkish subject area and two curriculum development specialists) took part in the analysis of the fifth sub-objective in line with the RBT, and a fourth researcher (a measurement and evaluation specialist) evaluated the results. For codes on which the researchers could not reach a consensus, the decision was made with a majority of votes. After the questions were coded in accordance with the RBT, two randomly selected questions from each cognitive level (10 questions in all) were presented to five experts from different fields of expertise (Turkish education, measurement and evaluation, and curriculum development) for their views. The experts' views were consistent with the decisions of the researchers, with the exception of the ninth question. Based on the evaluation of two experts for the ninth question, an additional explanation is given under Table 1.

## Results

### *Findings regarding the similarities of Turkish course exam papers*

The aim of the first sub-problem of the research was to analyse the similarity rates of the Turkish course exam papers examined within the scope of the study. The descriptive statistics regarding the similarity rates of the Turkish course exam papers examined within the scope of the study are as seen in Table 2.

**Table 2.** Descriptive Statistics for Similarity Rates of Turkish Course Exam Papers

Similarity Levels	Similarity Rate Ranges (%)	f	%
4th level	11-24	9	17.65
3rd level	25-49	19	37.25
2nd level	50-74	16	31.37
1st level	75-91	7	13.73
Total		51	100.00

When Table 2 is examined, it is seen that the Turkish course exam papers had similarity rates ranging from 11% to 91%. 17.65% of the exam papers examined had fourth level, 37.25% had third level, 31.37% had second level, and 13.73% had first level similarity rates. It is seen that among the exam papers examined, the most were in the third level similarity rate range. There were 19 exam papers in total at the third level similarity rate, and the exam papers at this level comprised 37.25% of all exam papers. This ranking is followed by the second level similarity rate with 16 exam papers and the first level similarity rate with 7 exam papers.

While determining the similarity rates, the websites utilised by Turkish teachers were also identified, and it was seen that a total of 174 websites were used for the 51 exam papers examined within the scope of the research. It was determined that the questions taken from these sites were transferred to the exam papers exactly as they were without making any changes. The sites with the highest usage rates for the exam papers were "dershanem.net (86%), turkceci.net (86%), bilimkoleji.com.tr (82%), egitimaski.com (72%), dilbilgisi.net (%52), harbiforum.net (43%), and simbioz.net (41%). The investigations showed that there were only 6 exam papers that benefited from the Education Information Network (EBA) of the Ministry of National Education. These exam papers, on the other hand, benefited from EBA at an average rate of 8.16%. The percentage of this rate among all exam papers is 0.96%.

### *Findings regarding the distribution of Turkish course exam papers in terms of page layout*

When the Turkish course exam papers were examined, it was concluded that 22 of them (43.14%) were in compliance with the page layout rules, while 29 of them (56.86%) were not. When the order of the items on the written papers was examined, it was concluded that 20 (39.22%) of them combined the questions in the same format, while 31 (60.78%) did not combine the questions in the same format, and that moreover, items with the same subject area/content were not used together.

**Findings regarding the item types and item writing criteria in the Turkish course exam papers**

When the question types in the 51 written exam papers for the Turkish course were examined, it was seen that 182 (24.36%) were prepared as open-ended, 388 (51.94%) were multiple-choice, 19 (2.54%) were true-false, 145 (19.41%) were short-answer and 13 (1.74%) were matching items. Moreover, it was found that 11 (21.57%) of the written papers included only multiple-choice items, nine (17.65%) contained only open-ended items (with short answers and long answers) and 31 (60.78%) were mixed tests (containing at least two different item types). However, considering the item types in the mixed tests, it was seen that multiple-choice item types were mostly used.

The distribution of teachers' item writing criteria for the fourth research question is tabulated separately for each item type. Accordingly, open-ended item writing criteria are given in Table 3, multiple-choice item writing criteria are shown in Table 4, short-answer item writing criteria are presented in Table 5, true-false item writing criteria are given in Table 6, and matching item writing criteria are shown in Table 7.

**Table 3.** Open-Ended Item Writing Criteria

Criteria	Yes		No	
	f	%	f	%
Are there any spelling, writing or expression errors in the item?	54	30	128	70
Is the item appropriate for the level of the students?	123	68	59	32
Is the item academically correct?	169	93	13	7
Is the item written clearly and understandably?	144	79	38	21
Is the item solved independently of the text/paragraph??	8	4	74	96
Is the item measure an important behaviour?	118	65	64	35
Is the item be answered independently of each other?	145	80	37	20

When Table 3 is examined, it is seen that there were spelling, writing or expression errors in 30% of the open-ended items used by Turkish teachers in their written exam papers, while there were academic errors in 7% of them. Moreover, it was found that 32% of the items were not appropriate for the students' level (below the students' level), while 35% did not measure an important behaviour or one that needed to be measured. It was observed that 79% of the items were written clearly and understandably, and that 80% of the items were answered independently of each other.

**Table 4.** Multiple-Choice Item Writing Criteria

Criteria	Yes		No	
	f	%	f	%
Is the item appropriate for the level of the students?	262	68	126	32
Is the item measure an important behaviour?	232	60	156	40
Is the item academically correct?	373	96	15	4
Are there any spelling, writing or expression errors in the item?	89	23	299	77
Are expressions that can give a clue to the correct answer avoided in the item root?	350	90	38	10
Is too much information included in the item root?	54	14	334	86
Are the item root and options written clearly and understandably?	335	86	53	14
If a negative expression/attention-grabbing expression is used in the item root, has the expression been stated conspicuously? (bold, underlined)	76	46	90	54
Are the distractors correct to a certain extent?	69	18	319	82
Has care been taken not to use contradictory expressions in the options?	340	88	48	12
Has care been taken to ensure that the lengths of the options are close to each other?	280	72	108	28
Have repeated statements been avoided in the options?	319	82	69	18
Are the options written independently of each other?	281	72	107	28

As seen in Table 4, it was found that approximately 23% of the multiple-choice items used by Turkish teachers in their written exam papers contained spelling, writing or expression errors, while 32% were not appropriate for the level of students. Moreover, it was seen that 54% of the negative/attention-grabbing expressions in the item root were not conspicuously stated. In addition, it was concluded that there were differences between the lengths of the options in 28% of the items, that the options were not written independently of each other in 28% of them, and that contradictory expressions were used in the options in 12% of them. It was seen that 86% of the items did not include too much information in the item root and that the item root and options were written clearly and understandably in 86% of them.

**Table 5.** Short-Answer Item Writing Criteria

Criteria	Yes		No	
	f	%	f	%
Is there too much information in the premise of the item?	14	10	131	90
Are there any spelling, writing or expression errors in the item?	32	22	113	78
Does the item have only one correct answer?	137	94	8	6
Is the item written clearly and understandably?	110	76	35	24
Does the item measure a single/important behaviour?	112	77	33	23
Is the item appropriate for the level of the students?	89	61	56	39
Is the item academically correct?	130	90	15	10

When Table 5 is examined, it was found that approximately 22% of the short-answer items used by Turkish teachers in their written papers contained spelling, writing or expression errors, and that 24% were not written clearly and understandably. It is seen that 10% of the items contained academic errors and that too much information was included in the root/premise of 10% of the items. Moreover, it was concluded that 61% of the items were appropriate for the level of the students and that 77% of them measured a single/important behaviour.

**Table 6.** True-False Item Writing Criteria

Criteria	Yes		No	
	f	%	f	%
Are there any writing or expression errors in the item?	4	21	15	79
Is the item absolutely true or false?	16	84	3	16
Do the true and false answers form a certain pattern?	8	42	11	58
Is the item written clearly and understandably?	15	79	4	21
Does the item measure a single/important behaviour?	6	32	13	68
Is the item appropriate for the level of the students?	12	63	7	37
Are related items given together?	9	47	10	53

According to Table 6, it was observed that 21% of the true-false items in the Turkish course exam papers included spelling, writing or expression errors, while 21% of them were not written clearly and understandably. It was seen that 42% of the items formed a pattern in their true-false answers, and that 53% of related items were not given together. In addition, it was concluded that only 32% of the items measured a single/important behaviour, while 63% were appropriate for the students' level.

**Table 7.** Matching Item Writing Criteria

Criteria	Yes		No	
	f	%	f	%
Are there any spelling, writing or expression errors in the premise and decision statements?	5	38	8	62
Is it clearly stated how the matching is to be done?	8	62	5	38

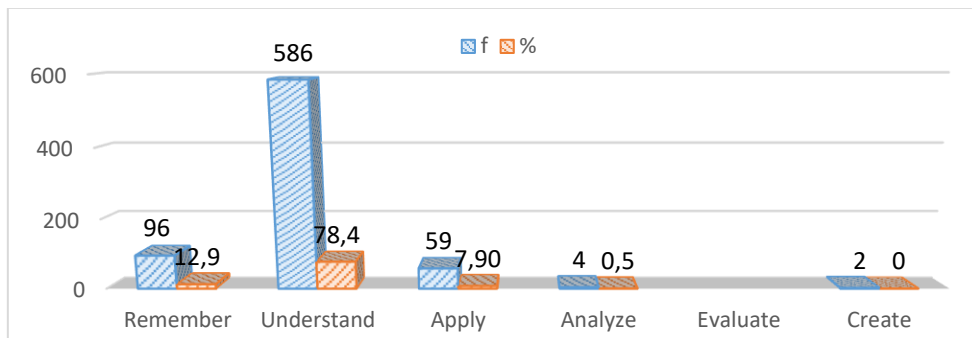
**Table 7.** Continued

Criteria	Yes		No	
	f	%	f	%
Has care been taken to have the same number of premise and decision statements?	3	23	10	77
Are the premise and decision statements written clearly and understandably?	10	77	3	23
Are the premise and decision statements appropriate for the students' level?	9	69	4	31
Has care been taken to ensure that the premise and decision statements measure the same characteristic?	7	54	6	46

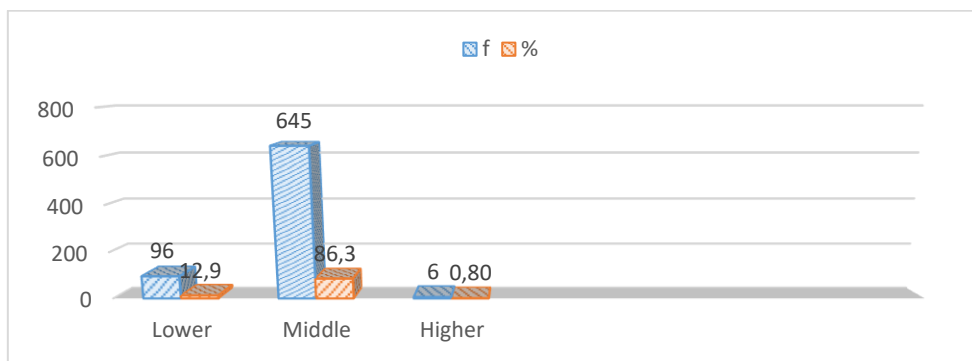
When Table 7 is examined, it is seen that 38% of the matching items in the Turkish course exam papers contained spelling, writing or expression errors, while 23% of them were not written clearly and understandably. It was found that 77% of the items did not have the same number of premise and decision statements, and that in 38% of them, it was not clearly stated how the matching was to be done. Moreover, when the premise and decision statements of the items were examined, it was concluded that the same characteristic/scope was measured in 54% of them, and that 69% of them were appropriate for the level of the students.

*Findings regarding the distribution of the questions in the Turkish Exam according to the cognitive process dimension and metacognitive level of the RBT*

The distribution of the questions in the Turkish Exam according to the cognitive process dimensions of the RBT is given in Figure 1, while information on the lower, middle and higher-order cognitive levels, in which the questions are located, is given in Figure 2.



**Figure 1.** Distribution of Questions According to Cognitive Process Dimensions of RBT



**Figure 2.** Distribution of Questions According to Cognitive Level

When Figure 1 is examined, it is seen that most of the questions were prepared at the *understand* level. Understand was followed by *remember*, *apply*, *analyze* and *create*, respectively. No questions were included at the *evaluate* level. When Figure 2 is examined, it is seen that the majority of the questions were aimed at the middle cognitive level. While questions for the lower level were then asked, the rate of questions for the higher-order level remained at the level of only eight per thousand.

## Discussion, Conclusion and Suggestions

In this study, Turkish course exam papers have been examined in terms of originality, page layout, item types, writing of items in accordance with their characteristics, and cognitive level. Within the scope of the first sub-problem of the research, the Turnitin program was used to examine the originality of the exam papers. The examination revealed that the exam papers had similarity rates ranging between 11% and 91% in terms of content. This finding of the study supports the finding by Akyol et al. (2013) that questions in exam books were used and that original questions prepared by teachers were rarely included. On the other hand, it is known that many domain experts emphasise that a measurement tool should not contain items and statements taken unchanged from a known source (Atılgan, 2006; Baykul, 2010; Doğan, 2006; Haladyna, 1994; Özçelik, 1992, 2013; Tekin, 2012).

Undoubtedly, the fact that the similarity rate of the exam papers is so high brings about many drawbacks in terms of measurement and evaluation requirements. The first of these drawbacks is that questions that can be freely accessed from the internet find their way into exam papers without undergoing any changes, a situation which increases the likelihood that students will encounter exam questions before the exam. This possibility raises the question that the measurement data do not reflect the correct information about the characteristic intended to be measured. The extent to which the functions of measurement and evaluation can be carried out under these conditions poses a serious question mark. Furthermore, as Atılgan (2006) states, even if a ready-made exam question has the quality of measuring a higher-order level behaviour, when there is a possibility that the question has been encountered before, only the recall behaviour of the student can be measured. In other words, since such an item does not measure the behaviour that it should measure, it cannot serve the purpose of measurement. It can be stated that the exam papers, which have a similarity rate of up to 91% as indicated by the research findings, can be suspected of not serving the purpose of measurement in this respect. Moreover, it should not be forgotten that an assessment method consisting of ready-made questions carries the risk of orienting the student to memorisation (Atılgan, 2006; Nitko & Brookhart, 2016). It is clear that the Turkish course, which has rich resources for developing students' higher-order thinking skills, should focus on ways for students to utilise their mental capacities in a more constructive way. Obtaining the questions ready-made from websites may lead students to orient towards these websites rather than the course. Therefore, while preparing for exams, it is possible that students will concentrate on internet data rather than on what is fostered by the course process.

The use of ready-made questions in the exam papers raises many questions, such as how compatible these questions are with the classroom context, how much they overlap with the outcomes in the curriculum, how much they reflect the points that the teacher emphasises in the teaching process, and to what extent the question difficulties are appropriate for the grade level. Indeed, Nitko and Brookhart (2016), who stated that ready-made questions may not match the content emphasised by teachers in their own lessons, state that these questions should be examined and checked by the teachers themselves in terms of their appropriateness for the course content and evaluation criteria. Therefore, instead of using these questions as they are, it would be a more correct way for teachers to adapt these questions in accordance with the measurement and evaluation principles, to use the ones that are appropriate for the outcomes they plan to measure, and to select the most appropriate question type for the purpose by knowing the strengths and weaknesses of the question types. In exams where such a process is not followed, the likelihood of making an erroneous assessment will be high (Nitko & Brookhart, 2016), reliability and validity will decrease, and decisions such as pass/fail given about the students according to the test results will be inaccurate. Research findings suggest that such problems exist, especially in exam papers having a high similarity rate.

As much as the similarity rates in the exam papers, the internet sites that Turkish teachers utilise when selecting exam questions are also an important topic of discussion. It was seen that a total of 174 internet sites were utilised for the exam papers examined within the scope of the research. None of the sites with the highest usage rate has the characteristic of being an official site. Furthermore, in none of the sites is it guaranteed that the contents have been prepared within the framework of the requirements

of educational sciences or that they have been audited by an expert. In addition, among all the papers, the percentage of exam papers benefiting from EBA was 0.96%. As well as the question of why the rate of teachers' use of original questions in exam papers is low, it is thought that the question of why teachers prefer other websites instead of EBA is one that needs to be answered. The research findings clearly reveal that teachers should be prevented from using questions published on websites, whose content has not undergone any inspection and which can be easily accessed by everyone, in their exams without making any changes. In this context, it is thought that software that makes it possible to examine the similarity rates in teachers' exams can be used in schools. Moreover, the development of platforms with encrypted access and questions that have been checked by experts for teachers to utilise in their exams can be suggested as a step that will improve the quality of Turkish exams.

In the examination made within the scope of the second sub-problem of the research, it was seen that 56.86% of the Turkish course exam papers did not have the page layout features that should be found in an exam paper. The format problems observed in the exam papers were that the letters had different sizes and fonts, and that appropriate space was not left between the item root and the options or items. Doğan (2006) stated that these criteria should be taken into account in tests in order to provide ease of perception and reading, and that otherwise, both validity and practicality problems will arise. In addition, it was concluded that in exam papers with mixed test characteristics, items of the same type were not found together and the items were randomly ordered. It is known that gathering items of the same type together is important for both teachers and students (McMillan, 2015). This feature, which is an essential for assessment tools, ensures that for the teacher, instructions are used less and scoring is easier, while for students, it prevents them from having to change their answering behaviours all the time. For this reason, while preparing the questions, it is necessary to pay attention to the page layout features and proper ordering of the item types, while failure to pay attention to these features brings out the validity problem in tests. Therefore, it was revealed that there was a validity problem in the majority of the exam papers examined within the scope of the research.

Another finding of the research is that the most used item type in the exam papers examined was the multiple-choice item type. This item type was followed by open-ended and short-answer item types. Another important finding is that in 21.57% of the exam papers, only the multiple-choice item type was used. This finding obtained in the study supports the findings by Aydın and Uçgun (2020), and Maden and Durukan (2009) that teachers mostly use the multiple-choice item type, while it differs from the finding by Benzer and Eldem (2013) that teachers mostly use the open-ended question technique. One of the reasons why the multiple-choice item type is preferred by teachers is that this item type can be scored easily and quickly in an objective way. The fact that teachers were preparing their eighth grade students for the high school entrance exam can be considered as another reason why this item type was preferred. Karakoç Öztürk and Altuntaş (2012) stated that teachers commonly use the multiple-choice test type because students are accustomed to this question type. However, the main objective in Turkish language education is for students to acquire the four basic skills, and it is clear that speaking, listening and writing skills, which are among these four basic skills, cannot be measured with multiple-choice items. Therefore, when the aim is to measure such skills as speaking, listening, writing, the effective use of language, problem solving, critical thinking, reflective thinking and analytical thinking etc., open-ended items and performance-based assessment practices should be used. On the other hand, in tests consisting of different item types, there will be advantageous or disadvantageous cases depending on the extent of the student's familiarity with the item format (such as multiple-choice items) (Hambleton, 2005). Although it was found in the research that multiple-choice items were mostly used in mixed tests, it was seen that mixed tests were preferred by teachers. It is seen that mixed tests are frequently used in large-scale tests such as PISA, TIMSS, etc. In in-class measurement and evaluation practices, too, teachers should carry out their measurement and evaluation practices by using item formats in a balanced way instead of using a single item type. Since different item types are used, mixed tests enable more valid and reliable measurements and decrease chance success by reducing errors caused by the effect of the item type.

When the results of compliance with the item writing criteria, which is another finding of the study, were evaluated as a whole, it was revealed that the teachers did not pay enough attention to the criteria in the items they used in the exam papers. For the criterion, "Are there any writing, expression or spelling mistakes?", it was observed that there were writing, expression and spelling errors in all item types and that some teachers did not consider this criterion. Considering that in Turkish, the use of the comma in the wrong place changes the meaning of the sentence, it can be concluded that this situation observed in the exam papers may cause students to misunderstand the item root or perceive the distractors as the correct answer. For this reason, teachers need to take this criterion into account in the items they write or those taken from different sources, and to make the necessary corrections before the item is added to the test. Otherwise, the reliability of the measurement results may decrease. In the examination made on the basis of the criterion, "If a negative expression/attention-grabbing expression is used in the item root, has the expression been stated conspicuously?", it was revealed that the teachers did not consider this criterion in more than half of the negative items they used. However, this situation may cause students to perceive the items in a positive way or to answer the item incorrectly even though they know it, which may lead to a decrease in the reliability, and therefore the validity, of the test. Considering the criterion, "Does the item measure an important behaviour?", it was revealed that important/critical behaviours were not measured in all item types (at rates ranging between 20% and 40%). Since it is impossible to measure all of the educational goals in the subject area with tests, critical behaviours for the subject area must be measured. For the criterion, "Do the true and false answers form a certain pattern?", it was seen that in approximately half of the items used (42%), the true answers to the items emerged in a systematic way. This means that in the event that students notice the pattern in the questions, they will be able to answer the item correctly without reading the question or knowing the answer. When these results are evaluated together, it is revealed that there is a need to focus in detail on pre-service and in-service teacher training processes to ensure that teacher candidates and teachers gain proficiency in item writing. In this respect, it can be recommended that the courses given at the undergraduate level be conducted under the supervision of a domain expert and a measurement and evaluation expert, and that practical studies be carried out. Working teachers can also be provided with in-service training on item writing, item writing criteria and item types. In these in-service training courses, it can be ensured that Turkish teachers are informed about what they will measure, and why, for what purpose and according to which criteria; and about what they should pay attention to in terms of validity and reliability in their exams; and that they have the chance to transform what they have learned into practice through workshops.

In the examination made within the scope of the fifth sub-problem of the research, it was concluded that approximately four-fifths of the questions used in the exam papers were prepared at the *understand* level of the RBT. While approximately one-fifth of the questions were aimed at the *remember* and *apply* levels of the RBT, no questions were prepared for the *evaluate* level, and very few questions were included for the *analyze* and *create* levels. The fact that the majority of the questions were prepared for the *understand* level is in parallel both with other research results for Turkish exams (Çintaş-Yıldız, 2015; Güfta & Zorbaz, 2008) and with results obtained for exams in different subjects (Himmah, Nayazik, & Setyawan, 2019; Virranmäki, Valta-Hulkkonen, & Pellikka, 2020). It can be thought that the reason why the exam questions were mostly aimed at the *understand* level is that this level includes more sub-skills than the other cognitive levels. While the other cognitive levels of the RBT include two or three cognitive sub-skills; the *understand* level includes seven sub-skills: interpreting, exemplifying, classifying, summarising, inferring, comparing and explaining. Frequently asking questions based on summarising, explaining and interpreting the given text in Turkish exams may be a reason why more questions are included for this level.

The Turkish course is an essential course in terms of fostering critical reading skills in students. Critical reading skills also require thinking skills such as analysing and evaluating (Ateş, 2013). However, among the questions examined in this study, only five per thousand were included in the *analyze* level, while no questions were included in the *evaluate* level. In the study conducted by Akyol (2001), it was determined that in the questions used in textbooks, questions that require thinking and



judgment such as evaluation were not included at all or were included very rarely. These results indicate that critical reading skill was not taken into account in the evaluation or that it was not known how to evaluate it, since although teachers included the phrase “evaluate” in the question root, they did not give an explanation about the internal and external criteria to be used in the evaluation. Based on this result, there is a need to carry out studies aimed at enabling teachers to write questions in the *analyze* level and especially, by including internal and external criteria, the *evaluate* level.

It was observed that among the questions examined, only three per thousand were at the *create* level. In the study by Aydın and Uçgun (2020), too, it was revealed that questions aimed at writing in exams were included at the knowledge level. This result shows that Turkish course exam questions focus on the “understand” dimension but do not contribute to “meaning-making”. Akyol (1997) argues that the skill of meaning-making should definitely be acquired by students and that this should be done in a planned way from primary school to university, since in order to enable skill development, practical instruction and assessment studies should be carried out. In Turkish lessons, writing activities for a particular type of text (fable, memoir, essay, etc.) are activities that can correspond to the *create* level. Among the eighth grade outcomes of the 2018 Turkish Curriculum are outcomes aimed at creative and critical writing (MoNE, 2019). However, the fact that teachers do not include questions for the *create* level is an obstacle to the development of writing skills. Since both writing questions in Turkish exams for the *create* level, which will contribute to the development of writing skills, and the evaluations at the *create* level require more time than for other types of questions, it can be recommended that such evaluations be made at separate lesson times. The unbalanced distribution of the questions in this study for the cognitive levels of the taxonomy reveals the idea that a taxonomy was not used in the exam preparation process or that the taxonomy was not used effectively by the teachers. A situation observed in different studies is that teachers did not benefit from a taxonomy during the preparation of exams (Ayvaci & Türkdoğan, 2010) and that they had difficulties in classifying the questions according to a taxonomy (Yusof & Hui, 2010). Since each level of the taxonomy necessitates the use of a different cognitive process, in instruction and assessment studies, it is important to include questions for each level of the taxonomy used (McMillan, 2015). Based on this importance, training should be given to teachers to ensure that they use the taxonomy effectively in the question preparation process.

Another striking result of the study is that among the examined questions, the rate of higher-order questions was only eight per thousand. Although the great majority of the questions were prepared for the middle cognitive level, the rate of questions for the lower cognitive level was about 13 percent. This result is consistent with the findings of studies (Güfta & Zorbaz, 2008; Kavruk & Çeçen, 2013; Maden & Durukan, 2009; Özen, 2020) examining Turkish exam questions. Three possible reasons can be suggested to explain why insufficient higher-order questions are included in Turkish exam papers. The fact that the questions were not written in accordance with the question-writing criteria, as seen in the fourth finding, can be considered as the first reason. The fact that in the examined questions, the answer to the question is given directly in the text or that the question root contains the answer prevents the activation of higher-order thinking skills. In the studies by Güfta and Zorbaz (2008), it was determined that the answers were included in the text that was given. As the second reason, it can be thought that teachers did not aim to write higher-order questions. As stated in the first finding, the high degree of similarity between the exams can be accepted as an indication that the questions were not prepared by the teachers and that the teachers did not aim to use higher-order questions in their exams. The fact that only 24 percent of the questions used in the exams (see the third finding) consisted of open-ended questions suitable for writing higher-order questions can be evaluated as the third reason. The fact that most of the questions were prepared as multiple-choice, true-false, short-answer and matching items creates limitations in the use of questions based on in-depth understanding and reasoning (McMillan, 2015) and may be insufficient for measuring advanced complex behaviours (Doğan, 2006). Another sub-reason that should not be overlooked is the fact that teachers often prepare open-ended questions for the lower and middle cognitive levels. In the study conducted by Özen (2020), it was seen that teachers prepared open-ended questions mostly as simple reading comprehension questions. However, an exam paper should include questions from different cognitive levels (Jones, Harland, Reid,

& Bartlett, 2009). While knowledge-level questions can be forgotten at a rate of 80-90%, higher-order level questions are remembered at a rate of 80-85% (Savage, 1998).

Having a balanced distribution aimed at different cognitive levels in exam papers and educating individuals with the higher-order thinking skills required for the 21st century requires following a systematic and planned instruction and measurement-evaluation process. For this reason, in their classroom teaching activities and their assessment practices aimed at monitoring student development, teachers should provide their students with situations and questions that require them to use these skills. After these thinking skills of the students are developed in the teaching process, questions measuring these skills should definitely be included in measurement practices carried out for level determination. For this reason, it may become obligatory to include higher-order questions in the exams implemented in schools. For the purpose of preparing exam questions, a text archive can be created that will make it possible to write higher-order questions, and this archive can be made available to teachers. Sample questions prepared by experts at different cognitive levels and including different thinking skills can be made available to teachers through EBA.

### **Limitations**

In this study, an attempt has been made to make comprehensive inferences about the quality and adequacies of Turkish exams by examining the exam questions in terms of originality, page layout, item types, writing questions in accordance with the item characteristics, and analysing them in terms of cognitive level. It is thought that in this respect, the study has contributing factors for the literature and for teachers. However, it also has various limitations. One of these limitations is that it was not known whether the students had previously come across the questions asked in the exams (during the course, in their homework, on the internet, etc.). For this reason, in all coding except for the *remember* level, it was assumed that the student first encountered these questions during the exam process. Another limitation is that the originality of the exam papers was limited only to the sources scanned by the Turnitin program. In addition, in terms of cognitive level, the use of only the Revised Bloom Taxonomy in the classification of the questions in the exam papers examined in the study, and the inclusion of exam papers belonging only to the first exam period in the sample, can be shown among the limitations of the study.

## References

- Acar-Erdol, T. (2020). Analysis of the questions in 11th Grade Philosophy Coursebook in terms of higher-order thinking skills. *Turkish Journal of Education*, 9(3), 222-245. doi:10.19128/turje.695928
- Aktaş, E. (2017). Öğretmen adaylarının farklı metin türlerine yönelik soru sorma becerilerinin yenilenmiş Bloom Taksonomisine göre değerlendirilmesi. *Electronic Turkish Studies*, 12(25), 99-118. doi:10.7827/TurkishStudies.12274
- Akyol, H. (1997). Okuma metinlerindeki soruların sınıflandırılması. *Education and Science*, 21(105), 10-17.
- Akyol, H. (2001). İlköğretim okulları 5. sınıf Türkçe kitaplarındaki okuma metinleriyle ilgili soruların analizi. *Educational Management in Theory and Practice*, 26(26), 169-178.
- Akyol, H., Yıldırım, K., Ateş, S., & Çetinkaya, Ç. (2013). Anlamaya yönelik nasıl sorular soruyoruz?. *Mersin University Journal of The Faculty of Education*, 9(1), 41-56.
- Anderson, L. W. (2005). Objectives, evaluation, and the improvement of education. *Studies in Educational Evaluation*, 31(2-3), 102-113. doi:10.1016/j.stueduc.2005.05.004
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R., Pintrich, P. R., ... Wittrock, M. C. (2001). A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives: Complete edition. New York, NY: Longman. Retrieved from <https://www.uky.edu/~rsand1/china2018/texts/Anderson-Krathwohl%20-%20A%20taxonomy%20for%20learning%20teaching%20and%20assessing.pdf>
- Anisah, N., Fitriati, S. W., & Rukmini, D. (2019). Teachers' questioning strategies to scaffold students' learning in reading. *English Education Journal*, 9(1), 128-143. doi:10.15294/EEJ.V9I1.28110
- Aslan, C. (2011). Soru sorma becerilerini geliştirmeye dönük öğretim uygulamalarının öğretmen adaylarının soru oluşturma becerilerine etkisi. *Education and Science*, 36(160), 236-249.
- Ateş, S. (2013). Eleştirel okuma ve bir beceri olarak öğretimi. *Turkish Journal of Education*, 2(3), 40-49. doi:10.19128/turje.181063
- Atılğan, H. (2006). Doğru yanlış testleri. In H. Atılğan (Ed.), *Eğitimde ölçme ve değerlendirme* (pp. 262-280). Ankara: Anı Publishing.
- Aydın, M., & Uçgun, D. (2020). Ortaokul Türkçe dersi sınav sorularının program'daki kazanımlara göre incelenmesi. *Ana Dili Eğitimi Dergisi*, 8(2), 343-356. doi:10.16916/aded.678478
- Ayvacı, H. Ş., & Türkdoğan, A. (2010). Yeniden yapılandırılan Bloom taksonomisine göre fen ve teknoloji dersi yazılı sorularının incelenmesi. *Journal of Turkish Science Education*, 7(1), 13-25.
- Baykul, Y. (2010). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması* (2nd ed.). Ankara: Pegem Akademi Publishing.
- Bekaroğlu, A. (2007). *İlköğretim 6. Sınıf Türkçe dersindeki yazılı sınav sorularının soru basamaklarına göre incelenmesi ve değerlendirilmesi (Kastamonu örneği)* (Unpublished master's thesis). Abant İzzet Baysal University, Bolu.
- Belet, Ş., & Sağlam, F. (2015). Türkçe dersinde kullanılan ölçme-değerlendirme yöntem ve tekniklerinin sınıf öğretmenlerine göre değerlendirilmesi. *Anadolu Journal of Educational Sciences International*, 5(1), 115-145.
- Benzer, A. (2019). Türkçe ders kitaplarının PISA okuma yeterlik düzeyleri ile imtihanı. *Research in Reading & Writing Instruction*, 7(2), 96-109. doi:10.35233/oyea.659740
- Benzer, A., & Eldem, E. (2013). Türkçe ve edebiyat öğretmenlerinin ölçme ve değerlendirme araçları hakkında bilgi düzeyleri. *Kastamonu Education Journal*, 21(2), 649-664.
- Beyhan, S. (2012). *Türkçe öğretmenlerinin ölçme ve değerlendirme uygulamalarına ilişkin görüşleri: Düzce ili örneği* (Unpublished master's thesis). Abant İzzet Baysal University, Bolu.
- Büyükalın, S. (2007). *Soru sorma sanatı*. Ankara: Nobel Publishing.

- Büyükalın-Filiz, S., & Yıldırım, N. (2019). Ortaokul Türkçe dersi öğretim programı kazanımlarının revize edilmiş bloom taksonomisine göre analizi. *Elementary Education Online*, 18(4), 1550-1573. doi:10.17051/ilkonline.2019.632521
- Büyüköztürk, Ş., Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2012). *Bilimsel araştırma yöntemleri* (11<sup>th</sup> ed.). Ankara: Pegem Akademi Publishing.
- Ceran, D., & Deniz, K. (2015). TEOG sınavı sorularının okuma becerisiyle çözülebilmek düzeyi. *Journal of Mother Tongue Education*, 3(2), 92-109. doi:10.16916/aded.62200
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks: Sage.
- Çintaş-Yıldız, D. (2015). Türkçe dersi sınav sorularının yeniden yapılandırılan Bloom taksonomisine göre analizi. *Gaziantep University Journal of Social Sciences*, 14(2), 479-497.
- Deniz, K., & Keray Dinçel, B. (2019). Türkçe öğretmenlerinin anlama becerilerinde ölçme ve değerlendirmeye yönelik görüşleri. *Journal of Mother Tongue Education*, 7(1), 28-64.
- Doğan, N. (2006). Çoktan seçmeli testler. In H. Atılğan (Ed.), *Eğitimde ölçme ve değerlendirme* (pp. 281-292). Ankara: Anı Publishing.
- Doğan, N. (2019a). Geleneksel ölçme ve değerlendirme teknikleri I: Yanıtı seçme gerektiren ölçme araçları. In N. Doğan (Ed.), *Eğitimde ölçme ve değerlendirme* (pp. 113-135). Ankara: Pegem Akademi Publishing.
- Doğan, N. (2019b). Geleneksel ölçme ve değerlendirme teknikleri II: Yanıtı yapılandırmayı gerektiren ölçme araçları. In N. Doğan (Ed.), *Eğitimde ölçme ve değerlendirme* (pp. 139-176). Ankara: Pegem Akademi Publishing.
- Döş, B., Bay, E., Aslansoy, C., Tiryaki, B., Çetin, N., & Duman, C. (2016). An analysis of teachers' questioning strategies. *Academic Journal*, 11(22), 2065-2078. doi:10.5897/ERR2016.3014
- Erdoğan, T. (2017). The view of primary school fourth grade students and teachers' questions about Turkish language lessons in the terms of the revised Bloom taxonomy. *Education and Science*, 42(192), 173-191. doi:10.15390/EB.2017.7407
- Esen, Y. D. (2019). *Öğretmenlerin ölçme ve değerlendirme alanına ilişkin yeterlilik algılarının ölçeklenmesi* (Unpublished master's thesis). Ankara University, Ankara.
- Gelbal, S., & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlilik algıları ve karşılaştıkları sorunlar. *Hacettepe University Journal of Education*, 33, 135-145.
- Göçer, A. (2005). *İlköğretim II. kademe Türkçe öğretiminde ölçme ve değerlendirme* (Unpublished doctoral dissertation). Atatürk University, Erzurum.
- Göçer, A. (2016). Lisansüstü eğitim gören Türkçe öğretmenlerinin yazılı sınav sorularının incelenmesi. *Uşak University Journal of Social Sciences*, 9(3), 23-37.
- Güfta, H., & Zorbaz, K. Z. (2008). İlköğretim ikinci kademe Türkçe dersi yazılı sınav sorularının düzeyleri üzerine bir değerlendirme. *Journal of the Cukurova University Institute of Social Sciences*, 17(2), 205-218.
- Güneyli, A., & Abbasoğlu, Ş. (2015). Türkçe dersi yazılı sınav sorularına ilişkin değerlendirme: Kıbrıs örneği. *Journal of Academic Studies*, 17(67), 53-76.
- Haladyna, T. M. (1994). *Developing and validating multiple-choice test items*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Boston, MA: Allyn and Bacon.
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. Merenda, & C. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Hillsdale, NJ: Lawrence Erlbaum Publishers.

- Himmah, W. I., Nayazik, A., & Setyawan, F. (2019). Revised Bloom's taxonomy to analyze the final mathematics examination problems in Junior High School. *Journal of Physics: Conference Series*, 1188, 012028. doi:10.1088/1742-6596/1188/1/012028
- Hogan, T. P. (2007). *Educational assessment*. USA: John Wiley & Sons, Inc.
- İnceçam, B., Demir, E., & Demir, E. (2018). Ortaokul öğretmenlerinin sınıf içi ölçme ve değerlendirmelerde yazılı yoklamalarda kullandıkları açık uçlu maddeleri hazırlama yeterlikleri. *Elementary Education Online*, 17(4) 1912-1927.
- Jones, K. O., Harland, J., Reid, J. M. V., & Bartlett, R. (2009). Relationship between examination questions and bloom's taxonomy. In *39th IEEE Frontiers in Education Conference* (pp. 1-6). doi:10.1109/FIE.2009.5350598
- Karacaoğlu, Ö. C. (2008). Öğretmenlerin yeterlik algıları. *Van YÜ Faculty of Education Journal*, 5(1), 70-97.
- Karadüz, A. (2009). Türkçe öğretmenlerinin ölçme ve değerlendirme uygulamalarının "yapılandırmacı öğrenme" kavramı bağlamında eleştirisi. *Uludağ University Faculty of Education Journal*, 22(1), 189-210.
- Karakoç Öztürk, B., & Altuntaş, İ. (2012). İlköğretim ikinci kademedeki konuşma eğitimine yönelik öğretmen görüşleri: Nitel bir çalışma. *Journal of Research in Education and Teaching*, 2(1), 342-356.
- Karatay, H., & Dilekçi, A. (2019). Türkçe öğretmenlerinin dil becerilerini ölçme ve değerlendirme yeterlikleri. *National Education Journal*, 48(1), 685-716.
- Kavruk, H., & Çeçen, M. A. (2013). Türkçe dersi yazılı sınav sorularının bilişsel alan basamakları açısından değerlendirilmesi. *Journal of Mother Tongue Education*, 1(4), 1-9. doi:10.16916/aded.15990
- Kubiszyn, T., & Borich, G. (2003). *Educational testing and measurement: Classroom application and practice* (7<sup>th</sup> ed.). New Jersey: John Wiley and Sons, Inc.
- Lipscomb, J. W. (1985). Is bloom's taxonomy better than intuitive judgement for classifying test questions?. *Education*, 106(1), 102-108.
- Maden, S., & Durukan, E. (2009). Türkçe dersi yazılı sorularının öğrenme alanlarına, soru tiplerine ve taksonomilerine göre değerlendirilmesi (Erzurum ili örneği). *Atatürk University Journal of Kazım Karabekir Education Faculty*, 19, 95-115.
- McMillan, J. H. (2015). *Sınıf içi değerlendirme, etkili ölçütlere dayalı etkili bir öğretim için ilke ve uygulamalar* (A. Arı, Trans.). Konya: Eğitim Publishing.
- Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference?. *American Secondary Edition*, 33(1), 49-64.
- Miller, M. D., Linn, R., & Gronlund, N. (2009). *Measurement and evaluation in teaching* (10<sup>th</sup> ed.). Upper Saddle River, NJ: Merrill, Prentice Hall.
- Ministry of National Education. (2006). Temel eğitime destek projesi "öğretmen eğitimi bileşeni" öğretmenlik mesleği genel yeterlikleri. Ankara: MEB. Retrieved from [http://web.deu.edu.tr/ilyas/ftp/ogretmenlik\\_meslegi\\_genel\\_yeterlikleri\\_2006.pdf](http://web.deu.edu.tr/ilyas/ftp/ogretmenlik_meslegi_genel_yeterlikleri_2006.pdf)
- Ministry of National Education. (2017). Öğretmenlik mesleği genel yeterlikleri. Retrieved from [https://oygm.meb.gov.tr/meb\\_iys\\_dosyalar/2017\\_12/11115355\\_YYRETMENLYK\\_MESLEYY\\_GENEL\\_YETERLYKLERY.pdf](https://oygm.meb.gov.tr/meb_iys_dosyalar/2017_12/11115355_YYRETMENLYK_MESLEYY_GENEL_YETERLYKLERY.pdf)
- Ministry of National Education. (2019). Türkçe dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar). Retrieved from <http://mufredat.meb.gov.tr/Dosyalar/20195716392253-02-T%C3%BCrk%C3%A7e%20%C3%96%C4%9Fretim%20Program%C4%B1%202019.pdf>
- Moodley, V. (2013). In-service teacher education: Asking questions for higher order thinking in visual literacy. *South African Journal of Education*, 33(2), 1-18.
- Nitko, A. J., & Brookhart, S. M. (2016). *Öğrencilerin eğitsel değerlendirmesi* (B. Bıçak, M. Bahar, & S. Özel, Trans.). Ankara: Nobel Publishing.

- Özçelik, D. A. (1992). *Eğitimde ölçme ve değerlendirme*. Ankara: ÖSYM Education Publishing.
- Özçelik, D. A. (2013). *Okullarda ölçme ve değerlendirme. Öğretmen el kitabı*. Ankara: Pegem Akademi Publishing.
- Özen, O. (2020). *Türkçe öğretmenlerinin açık uçlu soru hazırlama becerilerinin incelenmesi* (Unpublished master's thesis). Atatürk University, Erzurum.
- Popham, W. J. (2005). *Classroom assessment*. USA: Allyn and Bacon.
- Sallabaş, M. E., & Yılmaz, G. (2020). Türkçe ders kitabında bulunan metin altı sorularının yenilenmiş bloom taksonomisi'ne göre incelenmesi. *Journal of Mother Tongue Education*, 8(2), 586-596. doi:10.16916/aded.679933
- Savage, L. B. (1998). Eliciting critical thinking skills through questioning. *Clearing House*, 71(5), 291-293. doi:10.1080/00098659809602727
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309-327. doi:10.1177/0265532213480128
- Suartini, N. K. T., Wedhanti, N. K., & Suprianti, G. A. P. (2020). Teacher's questioning strategies in junior high school: A case study. *Jurnal Pendidikan Bahasa Inggris Undiksha*, 8(2), 97-103.
- Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1-14. doi:10.1186/s40468-019-0077-8
- Tekin, H. (2012). *Measurement and evaluation in education* (21<sup>st</sup> ed.). Ankara: Yargı Kitap ve Yayınevi.
- Thorndike, R. L. (1971). *Educational measurement* (2<sup>nd</sup> ed.). Washington, D.C.: American Council on Education.
- Tofade, T., Elsner, J., & Haines, S. T. (2013). Best practice strategies for effective use of questions as a teaching tool. *American Journal of Pharmaceutical Education*, 77(7), 1-9.
- Uymaz, M., & Çalışkan, H. (2019). Öğretmen yapımı sosyal bilgiler dersi sınav sorularının yenilenmiş Bloom taksonomisine göre incelenmesi. *Kastamonu Education Journal*, 27(1), 331-346. doi:10.24106/kefdergi.2637
- Ülger, Ü. (2003). *İlköğretim 6, 7, 8. sınıflarda Türkçe dersi yazılı sınav soruları üzerine bir değerlendirme* (Unpublished master's thesis). Gazi University, Ankara.
- Virranmäki, E., Valta-Hulkkonen, K., & Pellikka, A. (2020). Geography tests in the Finnish Matriculation Examination in paper and digital forms – An analysis of questions based on revised Bloom's taxonomy. *Studies in Educational Evaluation*, 66, 100896. doi:10.1016/j.stueduc.2020.100896
- Watkins, R., Meiers, M. W., & Visser, Y. L. (2012). *A guide to assessing needs*. Washington DC: International Bank for Reconstruction and Development.
- Wilens, W. (1991). *Questioning skills for teachers. What research says to the teacher* (3<sup>rd</sup> ed.). Washington, DC: National Education Association.
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162. doi:10.1016/j.tate.2016.05.010
- Yaşar, Ş., Gültekin, M., Türkkkan, B., Yıldız N., & Girmen, P. (2005). Yeni ilköğretim programlarının uygulanmasına ilişkin sınıf öğretmenlerinin hazırbulunuşluk düzeylerinin ve eğitim gereksinimlerinin belirlenmesi (Eskişehir ili örneği). In *Eğitimde Yansımalar: VIII. Yeni İlköğretim Programlarını Değerlendirme Sempozyumu bildiri kitabı* (pp. 51-63). Ankara: Sim Matbaası.
- Yıldız, M., Kanık Uysal, P., Bilge, H., Patricia Wolters, A., Saka, Y., Yıldırım, K., & Rasinski, T. (2019). Relationships between Turkish eighth-grade students' oral reading efficacy, reading comprehension and achievement scores on a high-stakes achievement test. *Reading Psychology*, 40(4), 329-349. doi:10.1080/02702711.2018.1555363
- Yusof, N., & Hui, C. J. (2010). Determination of Bloom's cognitive level of question items using artificial neural network. In *2010 10th International Conference on Intelligent Systems Design and Applications* (pp. 866-870). New Jersey: IEEE.