Examining Social Studies Course Questions in Terms of Question Type and Test Item Writing Techniques

Ahmet Çopur, Asst. Prof. Dr., Corresponding Author
Balıkesir University, Türkiye
ahmetcopur@balikesir.edu.tr
Orcid ID: 0000-0003-3837-7260

Hakan Önal, Prof. Dr.
Balıkesir University, Türkiye
onal@balikesir.edu.tr
Orcid ID: 0000-0001-8434-9880

Abstract

This study aimed to determine the question type of the written exam questions prepared by the social studies teachers and to examine the compatibility of the questions with the item writing techniques by using the document analysis method. Study data were collected from 20 social studies teachers and analyzed through the "Item Writing Techniques Checklist". Research findings revealed that social studies teachers preferred multiple-choice question types the most in written exam questions, followed by two-choice, fill-in-the-blank, matching and open-ended question types, respectively. In addition, findings showed that teachers were open to improvement reading criteria such as phrasing the stem of the question clearly, directly and simply, creating equivalent clauses in grammar, length, and complexity, distractor development, shaping items around a single proposition or idea, stating that each equivalent may not be used once, more than once, or not at all, including more responses than premises, not leaving spaces of the same width, and hinting at the correct answer. Based on these result, it I suggested to conduct applied studies to investigate the reasons why open-ended questions are rarely used although they are crucial in terms of developing students' high-level thinking skills as well as to determine the level of acquisition of the achievements in the analysis, synthesis and evaluation steps and to develop projects to ensure the professional development of teachers in terms of acquisition-based and skill-based item writing techniques.

Keywords: Document review, exam question, question type, written exam.

Introduction

Education, which exists in every moment of life formally or informally, is not only a phenomenon where information is given and received; it affects the development and shaping of values and skills, as well as some personality traits of a person. Therefore, the form, content and goals of education are in a constant change and transformation from past to present. The latest development in this regard is the goals regarding 21st Century skills, which are set not only for teachers but also for students. Within the scope of these skills, people in training and education are expected to go through a systematic education in which knowledge, skills, values and attitudes are acquired together. Content, effectiveness, skills and evaluation are the common denominators included in the general definitions of education as well as in education studies.

A major transformation was undertaken in the curriculum in 2005 in Türkiye in line with the 21st century skills. This transformation was not only related to changing the content in textbooks and curricula, it was also philosophical in nature and included the entire education and training process. With this transformation, the affective and psychomotor acquisitions in the curriculum decreased and the number of cognitive acquisitions increased (Ministry of National Education [MoNE], 2005). In the meantime, teachers were encouraged to use alternative assessment and evaluation methods in the measurement and evaluation process to effectively measure the cognitive domain. To do this, teachers should have the knowledge and skills to create measurement tools including different question types.

Although different types of questions are written by teachers, classical written and multiple-choice examination questions are still the most important measurement tools. Classical written examination exams are frequently preferred by teachers because they can easily provide connections between information, dilemma situations, supporting or rejecting ideas, using them in new situations, making analysis, generating new ideas and making evaluations. Since it is difficult to find another type of assessment or test that can combine so many purposes, written probe questions can generally be considered the best means of measuring knowledge (Bektaş & Kudubeş, 2014; Çopur & Önal, 2023). In

fact, although many alternative measurement and evaluation tools have been suggested to be used with the constructivist approach (Gelbal & Kelecioğlu, 2007) that has been implemented in Türkiye since 2005, evaluation of student success continues to be made with traditional measurement tools that focus on product rather than teaching.

Bahar (2001) states that all evaluations other than traditional evaluation, including multiple choice tests, can be included in the scope of alternative measurement and evaluation. Yıldız and Uyanık (2004), who tried to reveal the difference between traditional measurement methods and alternative measurement methods, emphasized that alternative measurement methods evaluate the process as well as the product. Similarly, Çepni and Ayvacı (2007) stated that the main purpose of using alternative measurement and evaluation methods is to ask students to perform a task appropriate to the relevant outcome and to ensure that success in the task is measured with valid and reliable measurement tools.

It is expected that the learning and teaching activities which are used in the transfer of the content consisting of the information that students should receive during the teaching process are both appropriate for the level of the students and have the characteristics specified in the program. Teachers are not only transmitters reflecting their own characters to the class, but also the sole controllers of the learning outcomes. They control the learning outcomes by using assessment. Assessment is done with different methods and periods to observe the latest state of learning. Although the exams, which generally follow specific processes, are seen as tools to rank student achievement based on scores, these evaluations are actually measurement tools that are important in reflecting not only the student outputs but also the outputs related to the teachers, the school and even the system (Ertürk, 1975; Black, 1995; Brookhart, 1997; Heverly & Fitt, 1994; MoNE, 2005; Tekin, 1994). Accurate evaluation provides reliable data to the system. These data are also a feedback for the development activities planned to be carried out in the system. Therefore, the place of evaluation in the review of the system is crucial.

The main purpose of the teacher, who carries out the teaching activities in a planned way, should be to raise students who can develop skills, think and solve problems. For this, teaching activities should increase the students' curiosity towards and lead them to learn physically and mentally (Yıldız, 2004, p. 2). A process that can focus on the body and mind must be prepared in order for learning to take place.

Teachers are the most important stakeholders in the preparation of classroom learning environments in the education and training process. Therefore, the ability of students to reach the goals specified in the education programs also depends on the teachers' ability to use the tools that will enable learning, to enrich the methods and techniques specified in the curriculum and to add heart and soul to them (Özden, 1998, p. 150).

Almost all countries have adopted a model for students going through teaching processes to raise them as individuals who can think critically, can question and have problem-solving skills. Questions are the key teaching strategy for developing and using all these concepts (Yıldız, 2004, p. 3) because when questions are asked in accordance with the principles and rules of the studies subject, students' thinking skills will be activated. Therefore, the questions that support teaching during the lesson, the study questions in the textbooks and the questions in the prepared activities should have the potential to define the problem and produce solutions by activating thinking. On the other hand, some studies on the cognitive levels of textbooks activities (Büken & Artvinli, 2020; Durukan & Demir, 2017; Yılmaz & Gazel, 2017; Yılmaz et al., 2021) emphasized the difficulty of obtaining high-level thinking skills

from students by using activities prepared with low-level thinking skills. Hence, questions should be prepared to activate the thoughts of those who are expected to solve them, not just the thoughts of the individuals who prepare them.

Assessment tools, the ultimate purpose of which is evaluation, should not be used only to measure academic achievement and rank students. Assessment and measurement should also aim to determine whether the mental skills of the students can be activated at the end of the training. The questions in the measurement tool prepared for this purpose should be developed in a way that includes the general characteristics of the students as well as the content and objectives of the education. Even when the characteristics of the venue where the education takes place and the content of the education are the same, not all students can think in the same way. The questions used in the measurement tool should be capable of measuring the learning at different levels for students with different learning speeds and ways of thinking caused by differences in physical and social environment, family structure and individual differences. According to Baysen (2006), who evaluated the questions teachers asked during the lesson, some of the questions do not go beyond the purpose of remembering scientific information, while some of the questions aim to make use of mental processes at a higher level than recall. The type of questions, the number of questions and the level of questions in the measurement tools have always attracted researchers' attention, probably originating from curiosity about the way teachers collect information from students. Kaya et al. (2002) stated that why and how the information will be used is directly related to the criterion in determining the question types in the measurement tools prepared by the teachers to collect information from their students. The answers to be received from the students can create a classical output like the evaluation results, as well as a guide to the teacher in the teaching process.

Teachers must obtain the right data in order to evaluate the learning outcomes correctly. Accurate data can be collected with the right measurement tools. Therefore, every exam may not mean that an accurate assessment has been made, since there may be problems directly related to the content such as question level and content validity, as well as incorrect use of questions arising from the person who prepared the question. While preparing exams, teachers generally use multiple-choice, open-ended, two-choice, matching, short-answer, and fill-in-the-blank question types. A measurement tool can be prepared by using only one of these item types, or by using more than one question type. According to Göçer (2015), in order for the exam to reach its real purpose, it must be written correctly in terms of form as well as content. For this reason, having the necessary skills to prepare questions is an important step in creating a measurement tool to ensure sound evaluation.

Questions prepared haphazardly without considering the rules may contain many typos in the writing of the question as well as the content and scope (Yavuz & Furkan, 2020). Each of the items in the exam prepared for assessment should be able to accurately measure the skills that are expected to be measured (Küçükahmet, 1999). For this, every item written in the test must be in accordance with the purpose and rules. Highlighting the constructivist approaches that have been implemented since 2005 has also diversified measurement tools. Teachers should be aware of the existence of these new measurement tools as well as how to prepare and implement them because the soundness of the evaluation depends on the accurate use of measurement tools. While discussing the benefits of using different question types, Çepni et al. (2003) emphasized that different question types also affect students' thinking levels. The questions have always attracted researchers' attention as they not only

affect students' thinking but also measure their attitudes towards the course and their academic achievements (Akpınar & Ergin, 2006). Linn and Gronlund (1995) defined question preparation as an art that requires certain knowledge and skills and emphasized that question preparation should be taught like all the other skills in all processes of education. In this regard, Karamustafaoğlu et al. 2003 mentioned the role of the teacher in increasing the quality of education and emphasized that teachers who can prepare quality questions contribute more in this process. Tekin (2004) stated that during the preparation of questions, rather than using the type of questions that should be selected for the relevant acquisitions, teachers generally prefer the question types they know best. In fact, many studies interpreted this finding as an error in item writing and suggested that teachers should be given in-service training on question preparation and identifying the required question type for relevant acquisitions (Anıl & Acar, 2008; Bakırcı et al. 2022;Gelbal & Kelecioğlu, 2007).

Examinations, which are generally regarded as measurement tools, can be considered in two main groups. The first of these are standardized exams, which consist of questions prepared in a systematic way, all stages of which are planned by experts, and are generally used in common exams for large audiences (Beydoğan, 1998). The second is the exams prepared by the teachers themselves to measure the learning outcomes of their classes by using their own potential. Balcı and Tekkaya (2000) stated that the exams defined in the second group and called teacher-made exams may not be very sound, and they emphasized that the teachers did not comply with the exam preparation processes and rules while preparing these exams which can be open-ended, multiple-choice, matching, true-false or fill-in-the-blank exams with long or short answers.

Like all curricula in Türkiye, the social studies curriculum consists of four basic elements: purpose, content, teaching-learning process, and evaluation (Demirel, 2015). The extent to which the first three of these can be achieved can be measured by evaluation. For this reason, proper measurement, and evaluation in social studies, as in every science, will ensure a correct understanding of both the process and the result.

Köken (2002) emphasized the importance of asking questions in social studies contcourses to measure whether students have acquired "knowledge", whether they understand it, whether they have made "analysis", at least not at the level of synthesis and evaluation, along with knowledge, interpretation, and application type questions. He stated that it should be done in a process that requires careful study. However, how well social studies teachers comply with the question preparation techniques and principles for the exams is a separate research topic. Bektaş and Kudubeş (2014) explain this by emphasizing that the exams held in schools are not for memorizing the details and testing their memorization, but for the purpose of using the information on the subjects, critical thinking and understanding.

According to the relevant literature (Çopur and Önal, 2023; Kılıç, 2010; Şanlı and Pınar, 2017; Uymaz and Çalışkan, 2019; Yılmaz and Gazel, 2017), while social studies teachers generally prefer multiple choice questions in their exams, they prefer open-ended questions less. Regarding the content validity of the questions prepared by teachers, although the questions have a slight relationship with the achievements in the program, it is among the common findings of many studies that the content validity is not fully complied with in terms of units. It is also emphasized in the studies in the literature that the cognitive levels of the questions asked by social studies teachers in exams mostly remain at the

remembering and understanding levels, and that the questions generally consist of questions corresponding to the lower cognitive level, while questions at the higher cognitive level are rarely asked.

In some studies (Gömleksiz and Bulut, 2007; Orbeyi, 2007; Sağlam-Arslan, Avcı and İyibil, 2008), it has been determined that although teachers are aware of the effectiveness of using constructivist assessment and evaluation tools, they prefer traditional assessment and evaluation approaches in practice. Undoubtedly, there may be more than one reason for this situation, but this study focused on the prepared questions rather than the selection of measurement and evaluation tools.

Traditional assessment and evaluation techniques generally consist of multiple choice, true-false, matching, fill-in-the-blank, short-answer and long-answer questions. On the other hand, alternative measurement and evaluation techniques consist of tools such as concept maps, diagnostic branched tree matching, word association, project, drama, show structured grid and product selection file (portfolio). Whichever method is used, teachers are expected to write questions in accordance with general rules when preparing questions.

Examination of the studies in literature investigating the questions prepared and implemented by teachers in Türkiye shows that almost all of them are related to the following topics: associating the questions with acquisitions (Kutlu et al., 2010; Aydın and Uçgun, 2020; Özgül and Polat, 2021) and determining the level of questions (Akpınar & Ergin, 2006; Baysen, 2006; Büken & Artvinli, 2020; Çepni et al., 2003; Karamustafaoğlu et al., 2003; Özgül & Polat, 2021). This study aimed to evaluate the question types used in the exams prepared by the social studies teachers (Balıkesir/Türkiye) themselves, and to explore these question types in terms of question and item writing rules. In this direction, the study sought answers to the following questions: "What is the distribution of written exam questions according to question types?" and "What is the compatibility of these question types with item writing techniques?"

Method

Research Design

Document analysis, one of the qualitative research methods, was used in the study. The collected data were analyzed with content analysis as well. Document review, which is defined as the sequence of operations that take place during the analysis and evaluation of electronic and/or printed materials (Bowen, 2009), consists of the stages of researching sources for a specific purpose, reading, and evaluating them (Karasar, 2005). This study was conducted based on document review, since it examined the compliance of written exam questions with item writing principles and included the above mentioned stages. In the study, the document review process was carried out with reference to the stages (Merriam, 2009) of finding the appropriate documents, checking the originality, determining the code-catalogue creation systematic and analyzing the data (content analysis). Based on the research problem, it was decided that the document analysis technique would be sufficient to collect data to shed light on the study problem.

Study Group

The study group was selected from among the social studies teachers employed in the districts of Balıkesir, Altıeylül and Karesi, by using convenience case sampling, one of the purposeful sampling techniques. Purposeful sampling technique was preferred due to the ease of identifying volunteer and

eligible participants (Patton, 2014). In this direction, the study was carried out with 20 volunteer social studies teachers.

Data Collection and Analysis

In the study, the data were examined through the "Item Writing Principles Checklist" (Annex-1), which was developed by the researcher using the literature (McMillan, 2015). Opinions were taken from Turkish Language and Literature (1) expert for the language control of the draft chart, and from measurement and evaluation (1) and social studies (1) expert for the content. Modifications were made in line with expert opinions. For example, "Was the problem expressed independently of the question stem?" and "Is the accuracy of the numerical answer confirmed?" and "Is the desired numerical unit clearly specified in the answer?" criteria were removed from the checklist in line with expert opinion. Finally, a 45-item item writing checklist was developed for multiple-choice, open-ended, two-choice, matching, short answer and completion (fill in the blanks) type questions. Table 1 presents the characteristics of the data collected within the scope of the research.

Table 1. *Characteristics of the analyzed data*

Exam Definitions	Exam Code	Number of Exams	Number of questions
5th Grade 1st Term 1st Exam	511	9	172
5th Grade 1st Term 2nd Exam	512	5	100
5th Grade 2nd Term 1st Exam	521	7	169
5th Grade 2nd Term 2nd Exam	522	6	128
6th Grade 1st Term 1st Exam	611	14	268
6th Grade 1st Term 2nd Exam	612	4	66
6th Grade 2nd Term 1st Exam	621	9	202
6th Grade 2nd Term 2nd Exam	622	6	127
7th Grade 1st Term 1st Exam	711	8	147
7th Grade 1st Term 2nd Exam	712	8	157
7th Grade 2nd Term 1st Exam	721	4	72
7th Grade 2nd Term 2nd Exam	722	5	108
Total		85	1716

The studies on item writing principles (McMillan, 2015) were examined for the analysis of the data obtained from the documents collected within the scope of the research, and 45 analysis criteria were determined. These criteria were designed to analyze multiple choice (10), open-ended (4), binary choice (10), matching (9), short answer (6) and completion (6) type questions. Each exam paper to be analyzed was coded and archived, and the analysis process was structured within this systematic framework. The data obtained within the scope of the research were analyzed twice by the researcher at different times and a part of the data (30%) was also analyzed by a social studies field expert. Finally, the analysis results of the researcher and the field expert were evaluated comparatively, and the agreements and disagreements were identified. After this stage, a common analysis table was created, and the process was terminated.

Validity, Reliability

In qualitative research, validity refers to the researcher's control for the accuracy of the findings, while reliability means that the researcher is consistent in his approach (Gibbs, 2007). Various strategies can be followed for the validity of qualitative research, such as member checking, triangulation, peer evaluation, and external auditor (Creswell, 2012). The reliability of the study was provided by following the steps of re-examination of the analyzed data at different times, continuous comparison of the data during the coding process, and cross-checking of the data by a different

researcher (Gibbs, 2007). Additionally, Miles and Huberman recommend that intercoder agreement be at least 80% for a high level of qualitative reliability (Miles & Huberman, 2016). In this direction, the data were checked again at different times to ensure the reliability of the study. In addition, cross-checking of the analyzed data was done by a field expert who was asked to analyze 30% of the data. The agreement (Miles and Huberman, 2016) as a result of the analysis of these two analyzes was calculated as 92%.

Ethical Permits of Research

In this study, all the rules specified to be followed within the scope of "Higher Education Institutions Scientific Research and Publication Ethics Directive" were complied with. None of the actions specified under the heading "Actions Contrary to Scientific Research and Publication Ethics", which is the second part of the directive, have been taken.

Ethics Committee Permission Information:

This study is within the scope of activities that do not require ethical permission.

Findings

What is the Distribution of Written Exam Questions by Question Types?

The question types in the written exams examined within the scope of the research were analyzed at the grade level, and the findings were given in Figure 1 and explained below. In order to facilitate the readability of the tables and figures, the item numbers for the criteria are included, and all of the criteria are presented in Annex 1.
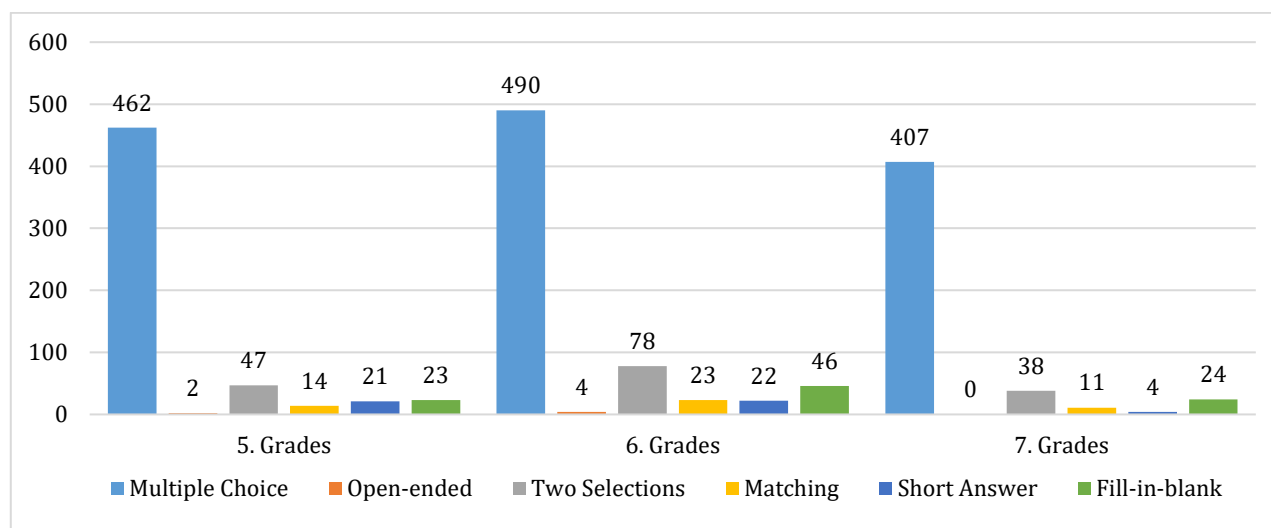

Figure 1. Distribution of written exam questions according to question types

Figure 1 shows that teachers mostly prefer multiple-choice question type at each grade level in written exams. This may be related to the fact that students will take general exams with multiple-choice questions and that such questions allow objective and rapid assessment. While the multiple-choice question type was followed by two-choice, fill-in-the-blank and matching, respectively, the least preferred question type was open-ended questions. The use of a total of six open-ended questions (two at the fifth-grade level and four at the sixth grade level) may be due to reasons such as the difficulty of evaluating open-ended questions and the difficulty students experienced in answering such questions.

Compliance of Written Exam Questions with Test Item Writing Techniques

Within the scope of this research question, the written exam questions were examined according to 45 criteria and three levels as Y (yes, appropriate), P (partially appropriate) and N (no, not appropriate) and provided under separate headings according to question types.

Compliance of Multiple-Choice Questions with Test Item Writing Techniques

Suitability of multiple-choice questions to question writing techniques was analyzed by using ten criteria. Figure 2 provides the findings.
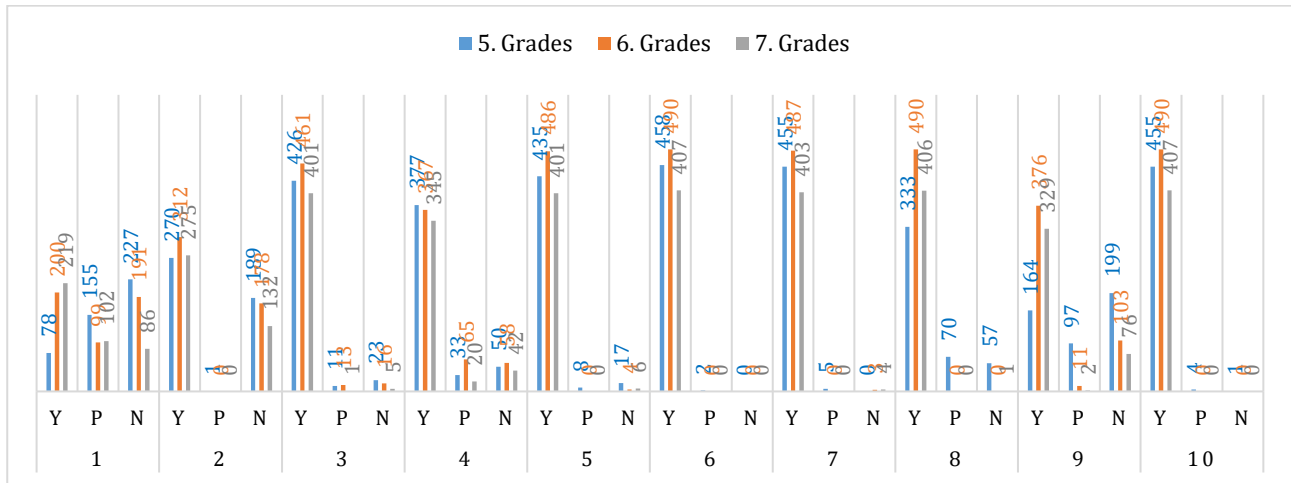

Figure 2. Suitability of multiple-choice questions to test item writing techniques.

According to the findings in Figure 2, it can be argued that the multiple-choice questions are open to development in terms of the following criteria: "Is the question stem phrased as clearly, directly and simply as possible? (1), "Is the use of negative expressions avoided in the stem of the question? (2)", "Are all options in line for grammatical structure, length, and complexity? (4)" and "Do distractors make sense for students who do not know the correct answer? (9)." In addition, although it was found that the questions were in compliance with the item writing techniques to a great extent in regards to the following criteria "Are the options short?" (6), "Are complex options avoided?" (7) and "Are the correct answers equally distributed among all options? (10)"; questions that did not comply with the criteria were also observed. Sample questions that formed the basis of these findings and that did not meet the relevant criteria are provided below.

Example 1. *Example of a question in which the stem of the question is not expressed clearly, directly and simply*

> ***Item 16.*** *"The Urartian King Menua opened the Shamran Canal to bring water to the city and irrigate the Tuspa Plain thanks to the power of God Haldi. Whoever says to someone that this canal was opened by someone esle, not by Menua, and wants to damage the canal; that person be destroyed God Haldi, God Teishaba and by all the gods"*
>
> *According to this text from the Urartian civilization, which of the following cannot be said about the Urartians?*
>
> *A) They believe in polytheistic religion.*
>
> *B) They lived on the banks of the Gediz River in Western Anatolia.*
>
> *C) Urartian King Menua tried to protect his work.*
>
> *D) They were engaged in agricultural activities.*

Example 2. *Example of a question in which the stem of the question is not expressed positively*

    Item 7. *Which of the following is not a result of the nomadic lifestyle of the early Turks?*

      A) The works of art are of portable material.

      B) Reaching advanced levels in agriculture

      C) Developing feelings of independence and freedom

      D) Their livelihood is based on animal husbandry and trade

Example 3. *Example of a question not appropriate in terms of grammar, structure and length*

    Item 10. *Which of the following is not one of the reasons why we cite references in our homework?*

      A) Respect for the labor of the owner

      B) Academic honesty

      C) The fact that people work on the same subject get more information on the subject

      D) Homework looks good

Example 4. *Example of a question without strong distractors*

    Item 12. *Which of the following is not one of the responsibilities of 6th grade student Metin?*

      A) Keeping the school clean

      B) Damaging the school property

      C) Being respectful to the staff

      D) Taking friends' belongings without permission

Examination of Example 1 in Figure 2 shows that the stem of the question was not in a simple structure. The stem of the question could be written more simply as, "Which of the following cannot be deduced about the Urartian from this text?" Example 2 includes a question in which the stem of the question was not expressed positively. The question stem can be expressed in a positive way as "Which is a result of the nomadic lifestyle of the early Turks?" The question sample (example 3) that is not suitable in terms of grammar, structure and length showed that the items differed from each other in terms of grammar and structure. These items can be written as"A) Respecting the work of the author, B) Paying attention to academic honesty, C) Ensuring that researchers on similar subjects gain more information, D) Enriching the work visually" to improve the item. Example 4, on the other hand, does not contain a strong distractor, allowing students who do not know the correct answer to find the correct answer easily.

Compliance of Open-Ended Questions with Test Item Writing Techniques

The suitability of open-ended questions to question writing techniques was analyzed by using four criteria. Figure 3 presents the findings obtained as a result of the analysis.
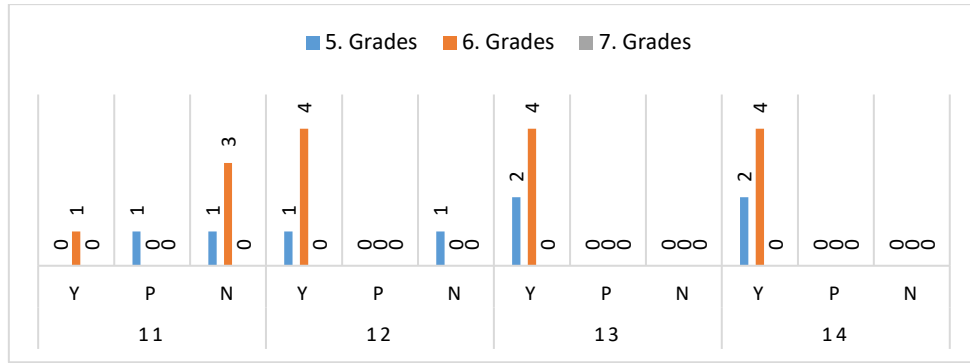
Figure 3. Suitability of open-ended questions to test item writing techniques.

The findings given in Figure 3 showed that the open-ended questions were open to improvement regarding the following criteria: "Were targeted reasoning skills measured? (11) and "Is the task clearly stated? (12)". Although, the questions were found to comply with the writing techniques to a great extent based on the criteria "Is the time given to answer the questions sufficient? (13) and "Was it possible to prevent students to choose from among various questions? (10)", it can be argued that the inclusion of very few questions in this question type constituted an obstacle to reach more comprehensive findings. Questions pertaining to this finding are presented in Examples 5 and 6.

Example 5. *Example of a question that are not suitable for measuring targeted reasoning skills*

    Item 21. *Define the concepts of congress, yuğ, kurgan, balbal and töre.*

Example 6. *Example of a question where the task is not clearly stated*

    Item 15. *What will be the negative consequences if we do not fulfill our responsibilities? Explain with an example.*

Examination of Example 5 in Figure 3 shows that the open-ended question was not prepared to measure the targeted reasoning skill. While Example 5 should have been prepared at the analysis stage within the scope of the learning outcome of "SS.6.2.1. The student makes inferences about the geographical, political, economic and cultural characteristics of the first Turkish states established in Central Asia", the question was prepared at the knowledge level. This question can be written as "What inferences can be made about the understanding of state administration, based on the fact that there was an assembly called a congress in the first Turkish states?" to measure the desired acquisition. the achievement: Example 6 is an example of a question in which the task is not clearly stated. By specifying the area of responsibility for this question, it can be written as "What are the consequences if we fulfill our responsibilities as a student/child/citizen? Explain with an example", to state what is expected from the student more clearly.

Compliance of Binary Choice Questions with Test Item Writing Techniques

The suitability of binary choice questions to question writing techniques was analyzed by using ten criteria. Figure 4 presents the findings obtained as a result of the analysis.
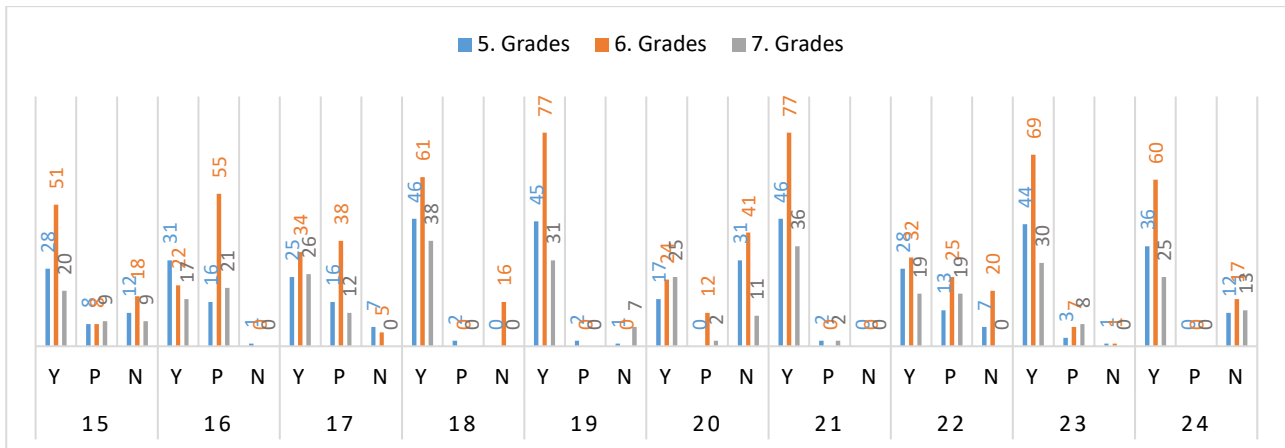
Figure 4. Suitability of binary choice questions to test item writing techniques.

According to the findings presented in Figure 4, binary choice questions may be open to improvement regarding the following criteria: "Does the item contain a single proposition or idea? (15), "Is there a pattern of correct answers? (20)" and "Is the item written in a way that does not try to trick students? (22)". Also, regarding "Is the item positively worded? (18), "Is the length of both statements in the article approximately the same? (19) and "Are clear terms used? (21)" criteria, it can be stated that the questions generally conform to item writing techniques. However, as can be seen in examples 7 and 8, there are also questions that do not meet the relevant criteria.

Example 7. *Example of a question that contains one idea or proposition and attempts to trick the student*

> **Item** *(T)/F 1. Empathizing and respecting differences cause prejudices*

Example 8. *Example of a question prepared in a specific model*

> *(F) Mesopotamia is the region between the Seyhan and Ceyhan Rivers.*

> *(T) The fact that the belongings of the deceased are placed in the graves in the Urartians shows that there is a belief in life after death.*

> *(F) In Mesopotamia, the buildings were made of stone, which enabled all the artifacts to reach the present day.*

> *(T) The Code of Hammurabi was prepared according to talion.*

> *(F) The Assyrians were engaged in maritime trade since they were established by the sea.*

> *(T) With the invention of writing, information was easily passed on to future generations.*

> *(F)Ziggurats were used for worship only.*

> *(T) The Hanging Gardens are among the works of Babylon.*

> *(F) Hittites are the Anatolian civilization founded in the city of Ephesus near Ankara.*

> *(T) The existence of a free thought environment in the Ions has paved the way for scientific studies.*

Examination of Example 7 in Figure 4 shows that the binary choice question contains more than one idea or proposition. This question can be written as, "Empathy prevents us from forming prejudices against different cultures" to include a single idea or proposition. Example 8, there is an example of a

question that increases the chance factor since the correct answers are prepared to follow a pattern. This situation can be eliminated by changing the places of the items in this question.

Compliance of Matching Questions with Test Item Writing Techniques

The suitability of matching questions to question writing techniques was analyzed by using nine criteria. Figure 5 presents the findings obtained as a result of the analysis.
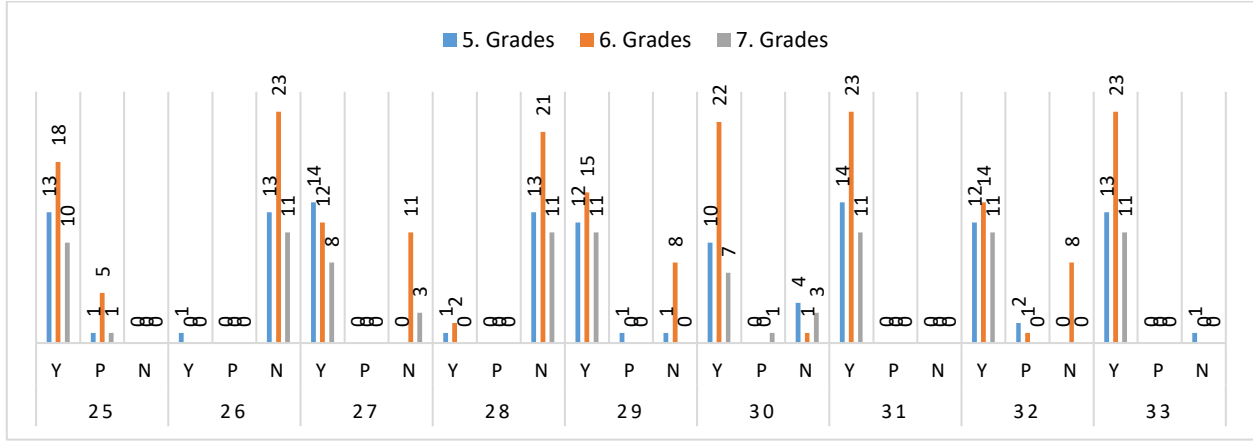

Figure 5. Suitability of matching questions to test item writing techniques.

Findings in Figure 5 showed that matching questions were open to improvement regarding the following criteria: ; "Is it clear that each option may be used once, more than once, or not at all? (26) and "Are there more responses than premises? (28)". Also, it can be stated that the questions largely conform to the writing techniques regarding "Is it clear where and how students will place the answers? (25), "Is there one possible response for each premise? (31) and "Are the premises and their counterparts on the same page? (33)" criteria. However, there are also questions that do not meet the specified criteria some of which are mentioned in Examples 9 and 10.

Example 9. Example of a question in which it is not stated that each answer can be used one or more times or not at all

Hz Ebubekir	Hz Ömer	Bilge Kağan	Kaşgarlı Mahmut	Balbal
Hz. Muhammed	Hakan Kağan	Mete Han	Gazneli Mahmut	Congress

Item 4. Match the people and concepts in the table above with the information in the table below.

	He was born in Mecca in 571 and invited people to Islam.
	He gave the Kök Turkish State its heyday and erected Orkhon inscriptions.
	He wrote the book "Divan-ı Lügati Türk" in order to teach Turkish to Arabs.
	He organized 17 expeditions to India and ensured the spread of Islam.
	In Central Asia, it was used to meet under the chairmanship of the khan to discuss important state affairs.
	They are statuettes erected on the grave of the deceased in Central Asian Turks.
	He made the Asian Hun State experience its heyday..
	He was the first caliph. During this period, the Quran was made into a book.
	It is the name given to the rulers in the Turkish states established in Central Asia.
	During his reign, Islam spread out of the Arabian peninsula and the country was divided into provinces.

Example 10. Example of a question in which the number of responses are not more than the premises

C. Read the explanations in the table below and write in the box which social science branch it is opposite by selecting it from the box below.

Geography / Anthropology / Sociology / Psychology / Archeology	What science am I?
I am a social science trying to make generalizations about people and their behavior and to understand human diversity.	
I am a science that studies the relations between man and the environment, the earth.	
I am a science that studies the objects from the ancient period by excavations.	
I am a science that studies the individual, groups, institutions, events and the relationships and interactions between them in society.	
I am a science that studies human emotions, thoughts and behaviours.	

Examination of Example 9 in Figure 5 shows it is not stated in the question that each answer can be used one or more times or not at all. The question can be developed by adding an explanation in the form such as "Each answer to this question may be used one or more times or not at all." In Example 10, there is an example of a question that increases the chance factor as the correct answers are in a pattern. This situation can be eliminated by changing the places of the items in this question.

Compliance of Short Answer Questions with Test Item Writing Techniques

The suitability of short answer questions to question writing techniques was analyzed by using nine criteria. Figure 6 presents the findings obtained as a result of the analysis.
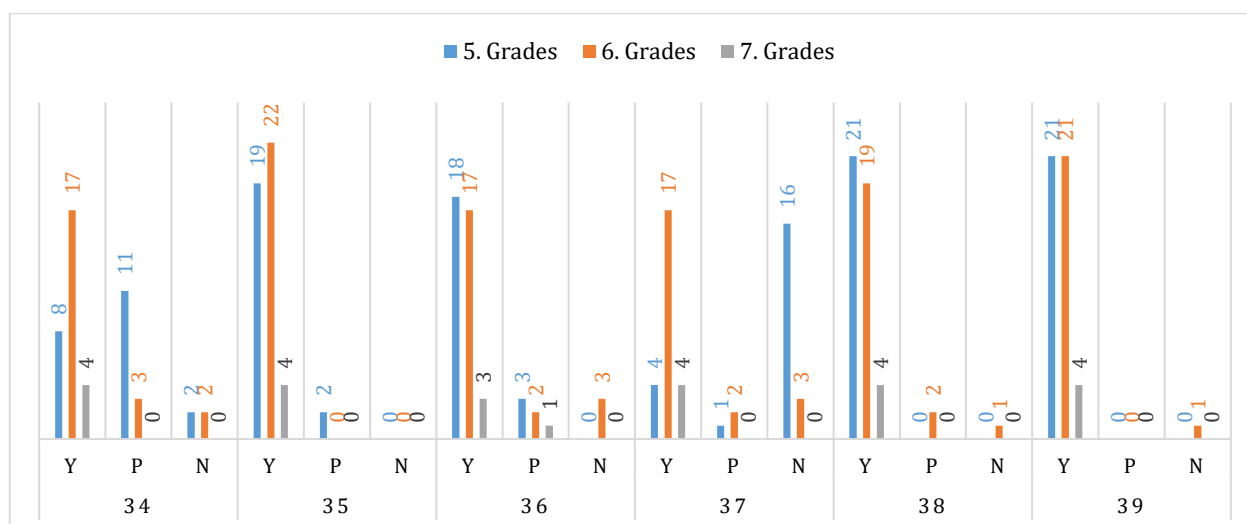

Figure 6. Suitability of short answer questions to writing techniques

Findings related to short answer questions in Figure 6 showed that they can be improved regarding the following criteria: "Is the question structured to have a single correct answer? (34) and "Was the question written as concisely as possible? (37)". In addition, while the criteria such as "Was it avoided to directly take questions from the student book? (35) and "Are the words used in the question difficult to understand for students? (39)," were generally complied with the writing techniques, the criteria number 37 "Was the question written as short and concisely as possible?" was found to be problematic especially for 5th graders. Examples of questions about to this finding are given in Examples 11 and 12.

Example 11. Example of a question not structured to have a single correct answer

Item: Give one example each about our roles in the following groups.

The group We Are Members Of	Our Role
5A Classroom	
Travel Tourism Club	
Family	
Basketball Team	
Photography Course	

Example 12. Example of question not worded shortly and concisely

Item: Examine the images below and write down which economic activity field they are related to.

Example 11 shows that that the question is structured to have more than one correct answer. This question could have been asked about roles in this group by specifying one of the groups students are a member of. For example, this question can be improved as "Give examples of our roles as students". Example 12 is a question that is not written concisely. In this example, the question stem is can be structured in a more concise and concise form as "Write the economic activities about following images."

Compliance of Fill-in-the-Blank (Completion) Questions with Test Item Writing Techniques

The suitability of fill-in-the-blank questions to question writing techniques was analyzed by using six criteria. Figure 7 presents the findings obtained as a result of the analysis.
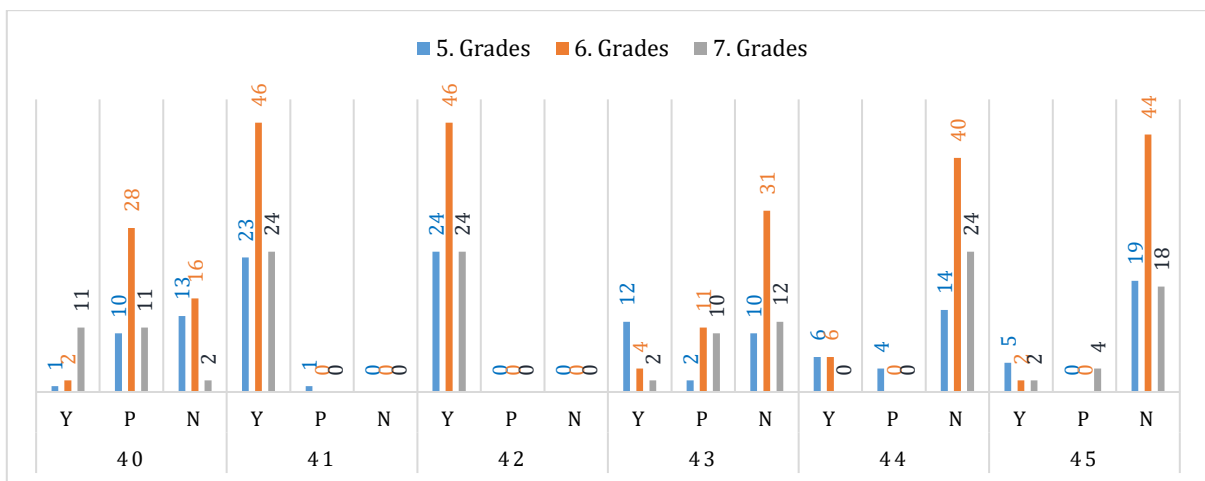

Figure 7. Compatibility of fill-in-the-blank (completion) questions with test item writing technique

Examination of the findings in Figure 7 showed that the Fill-in-the-Blank (completion) questions could be improved regarding the following criteria: "Is the use of verbatim language from teaching materials avoided? (40)", "Is the blank left at the end of the sentence? (43)", "Are the blanks all the same width? (44)" and "Is giving clues about the correct answer avoided? (45). In addition, it can be argued that the questions are not very suitable for writing techniques regarding the "Does the item evaluate knowledge? (41) and "Is a short and concise answer requested (42) criteria as well. Various negative aspects were observed especially regarding criteria 40, 43, 44 and 45. Examples of questions that do not meet the required criteria are presented in Example 13, 14, 15 and 16.

Example 13. *Example of a question using verbatim language from teaching materials*

Item 1. *In the blanks below, write the appropriate word from the given words. (10 points)*

Plain / Population / Maquis / Sketch / Steppe / Plateau / Natural Disaster / Map

Large fertile plains that are low compared to their surroundings are called ………………

The transfer of all or a part of the earth to a flat area by reducing it within a certain scale is called ………………

The number of people living in a certain area is called…………..

Events that occur largely or completely outside the control of people and cause loss of life and property are called ………………

The basic vegetation of the Mediterranean climate consisting of short grass communities is …………………

Example 14. *Question example where the blank is not at the end of the question*

Item: *B) Fill in the blanks in the following expressions with the words given.*

Tourism/Wood/Plato/UNICEF/SocialEnvironment/Production/Distribution/Bilindi/Consmer/

Respect for Labor

1) …….. is an organization established under the United Nations for the protection of children's rights.

2) The important elements of our ……… are holidays, folk dances and cuisine

3) ……is a wide plain deeply pierced by the rivers and higher than its surroundings

4)……is a place that does not exist in reality, but keeps people together through the internet.

5) ………… is the stage of delivering the produced goods and services to the consumers.

6) Due to the forest cover in the Black Sea climate, …………..houses are common.

7) ………. Is shown by using footnotes and bibliography in a scientific research ……….

8) …………… has developed in places with long summers, coastal, historical and natural beauties.

9)……………… is the person who makes purchases according to his needs and receives the receipt and invoice of the product while shopping.

10) ……..is the creation of a product to meet people's needs.

Example 15. *Example of a question where all the blanks are not the same width*

> ***Item 2) Place the following concepts in the blanks in a way that is appropriate and meaningful.***
>
> ***Tuspa/Writing/Assyrians/Anal/Money/Ziggurats***
>
> *1) The capital of the Urartians, founded in Eastern Anatolia, is ....*
>
> *2) With the invention of .... the ages of history began.*
>
> *3)........................ taught Anatolia to write thanks to trade.*
>
> *4) The Hittites wrote down the good and bad things they experienced for a year in the annals called ....*
>
> *5) With the invention of ...., the barter in trade was ended.*
>
> *6) The temples belonging to the Sumerians were called ....*

Example 16. Example of a question hinting the correct answer

> ***A-In the blanks below, write the appropriate word from the given words. (20 Points)***
>
> ***Balbal / Belief of Kut / Inscription / Yuğ / Afterlife / Faith / Kurgan / Sagu***
>
> *1) The fact that Turks bury the dead in the graves, along with their loved ones, shows that they believe in ....*
>
> *2) In the first Turkic states of Central Asia, the funeral and mourning ceremony held after the dead is called ....*
>
> *3) The belief in Turks that the authority to rule the state is given by the god is........*
>
> *4) In Central Asia, the big graves where Turks buried their dead with their belongings are called.......*
>
> *5 In Central Asia, the statues erected on the heads of the graves by the Turks as much as the number of enemies they killed are called.......*

Examination of Example 13 in Figure 7 shows that the teaching materials were quoted using verbatim in the question. This question may be improved by avoiding the use of verbatim langıage in the question stem such as, "The dominant vegetation reflecting the climatic characteristics in areas where the Mediterranean climate is observed ....?". In Example 14, the question stems can be restructured so that the blanks are at the end such as "The organization operating under the United Nations for the protection of children's rights is ....". The question example with unequal gap widths (example 15) can be improved by restructuring the gaps to be equal. Example 16 is written in a structure that gives clues about the correct answer. Questions can be improved by rewording such as "The reason why Turks bury their loved ones and horses in the graves along with the dead is ....".

Discussion and Conclusion

Research findings reveal that the most preferred question type by teachers is multiple choice questions followed by binary choice, fill-in-the-blank and matching, respectively, while the least preferred type is the open-ended questions.

Similar to the results of this research, Şanlı and Pınar (2017) found in their study examining the social studies course exam questions that while multiple choice questions were the most preferred question type by the teachers, open-ended questions were one of the question types preferred the least.

It is noteworthy that teachers do not use open-ended questions very much in written exams. Despite the difficulty of making a judgment based on the few open-ended questions examined in this study, it was observed that these questions were far from measuring the targeted reasoning skills and the task was not clearly stated in the questions.

While these findings are in line with the findings that multiple-choice questions are the most preferred question type by social studies (Kılıç, 2010; Şanlı and Pınar, 2017; Uymaz and Çalışkan, 2019), history (Danışmaz, 2019) and Turkish (Aydın and Uçgun, 2020) teachers, this contradicts the studies that concluded that teachers mostly ask questions with short-answers (Safran, 2002) and open-ended (Çolak, 2008) questions in exams. This may be due to the differentiation of question type selection by social studies and history teachers depending on the developmental characteristics of the students, as well as the fact that teachers change their question type preferences over time. It can be stated that use of multiple choice questions by social studies teachers may be related to finding this question type to be more comprehensive to ask questions in their field, ease and objectivity in evaluation (McMillan, 2015). In addition, the use of this question type in central selection and placement exams can be considered as another reason why teachers prefer multiple choice questions more. Less preference for open-ended questions may be due to the difficulty of students in answering such questions and the difficulty of evaluating them (McMillan, 2015). In addition, the fact that teachers generally prefer the question types they know best (Tekin, 2004) rather than the most appropriate question type to measure the relevant outcome (Tekin, 2004) may be another reason for this situation.

The obtained findings pointed out that teachers need to be improved in terms of phrasing the question stem clearly, directly and simply, writing the question items by using positive expressions, creating equivalent items in terms of grammar and length, and writing distractor items. In support of the findings, Downing (2005) conducted a study on the examination of 219 items in four tests and determined that the tests violated the article writing principle by 35% to 65% and that the most common violation in two tests was negative writing of the item. Tarrant et al. (2006) also determined that the correct answer being longer than the distractors was among the most common violations. These findings also coincide with the study findings. Aybek, Kartal and Yildirim (2022) found that the main idea is clearly and clearly expressed at the root of the article in the vast majority of the articles in both the field of science and social studies in their studies, where they examined their questions in the science and social studies course in terms of compliance with the principles of article writing. It can be said that these determinations partially coincide with the findings of the current research on expressing the question root in a simple way. In addition, in the writing of binary option questions, teachers need to be improved in using criteria such as shaping the items around a single proposition or idea and not selecting the correct answers according to a pattern. Aybek et al (2022), who reached similar findings, found that there are substances that require a unified response in their studies. In writing of matching type questions, teachers were found to be open to improvement regarding the criteria of stating that each response may not be used one or more times or not at all and including more responses than the premises. In addition, there are problems in criteria in the completion (fill- in-the-blank) questions; such as the use of verbatim language from teaching materials, positioning and structuring of the blanks, and giving clues about the correct answer.

Each of the items in the exam prepared for assessment should be able to accurately measure the skills that are expected to be measured (Küçükahmet, 1999). For this, every written item must be

accurate in terms of both the form and the content (Göçer, 2015). Questions prepared hastily without considering the rules may contain many typos in terms of question writing technique as well as content and scope (Yavuz & Furkan, 2020). In fact, this situation is interpreted in many studies as errors showing the need that teachers should be given in-service training on identifying question types and preparing questions (Gelbal & Kelecioğlu, 2007; Anıl & Acar, 2008; Bakırcı, Özmantar, & Bozkurt, 2022).

Since using different question types affects students' thinking levels (Çepni, Özsevgeç, & Gökdere, 2003), it can be argued that preparing written exams in accordance with test item writing techniques is crucial. Reaching findings in support of this issue, Karamustafaoğlu et al. (2003) cited the role of the teacher in increasing the quality of education and emphasized that teachers who can prepare quality questions contribute more in this process. In this regard, Balcı and Tekkaya (2000) stated that the evaluation made by teacher-made exams is not very sound because teachers do not comply with the exam preparation processes and rules. It can be said that the findings obtained from the studies mentioned so far support the current research findings bypointing out the importance of preparing questions in accordance with the test item writing techniques.

Recommendations

Based on the findings obtained from the research, the following suggestions are provided;

- Research results showed that multiple choice questions are preferred the most by teachers, followed by binary choice, fill-in-the-blank and matching question types, respectively. When determining the question type, teachers tend to prefer the question types they know the best (Tekin, 2004) rather than selecting the most appropriate question type to measure the relevant outcome (Tekin, 2004). Hence, it is suggested to provide trainings to increase teachers' proficiency regarding question types and test item writing principles,

- According to the findings, the least preferred question type by the teachers is the open-ended questions. It is suggested to explore the reasons and find solutions to the low level use of open-ended questions, which are very important to develop students' high-level thinking skills and to determine the level of acquisitions in the analysis, synthesis and evaluation stages.

- The findings suggest that teachers are open to improvement regarding the criteria such as phrasing the question stem clearly, directly and simply, writing the question items by using positive expressions, creating equivalent items in terms of grammar and length, writing distractor items, shaping the items around a single proposition or idea, not selecting the correct answers according to a pattern, stating that each response may not be used one or more times or not at all, including more responses than the premises, positioning and structuring of the blanks, and giving clues about the correct answer. Based on these results, suggestions were made to make projects to ensure the professional development of teachers in terms of question preparation techniques.

References

Akpınar, E., & Ergin, Ö. (2006). Fen bilgisi öğretmenlerinin yazılı sınav sorularının değerlendirilmesi [Evaluation of written exam questions of science teachers]. *Milli Eğitim Dergisi, 172*(Güz), 225-231.

Anıl, D., & Acar, M. (2008). Sınıf öğretmenlerinin ölçme değerlendirme sürecinde karşılaştıkları sorunlara ilişkin görüşleri [Elementary school theachers' views on issues they experience through measurement and evaluation processes]. *Van Yüzüncü Yıl University Journal of Education, 5*(2), 44-61.

Aybek, E. C., Kartal, S. K., & Yıldırım, Ö. (2022). Millî Eğitim Bakanlığı tarafından yayınlanan test maddelerinin madde yazım ilkelerine uygunluğunun incelenmesi [Investigation of the test items published by Ministry of National Education according to the ıtem writing guidelines]. *Bogazici University Journal of Education*, *39*, 357-374.

Aydın, M., & Uçgun, D. (2020). Ortaokul Türkçe dersi sınav sorularının Program'daki kazanımlara göre incelenmesi [An analysis of secondary school Turkish language course exam questions in the sense of curriculum objectives]. *Journal of Mother Tongue Education, 8*(2), 343-356.

Bahar, M. (2001). Çoktan seçmeli testlere eleştirel bir yaklaşım ve alternatif metotlar [A critical approach to multiple choice tests and alternative methods]. *Kuram ve Uygulamada Eğitim Bilimleri, 1*, 23-38.

Bakırcı, G., Özmantar, M. F., & Bozkurt, A. (2022). Ortaokul matematik öğretmenlerinin veri işleme öğrenme alanına dair yazılı sınav soruları üzerine inceleme [Investigation of the secondary school mathematics teachers written exam questions on data handling]. *E-International Journal of Educational Research, 13*(3), 1-17.

Balcı, E., & Tekkaya, C. (2000). Ölçme ve değerlendirme tekniklerine yönelik bir ölçeğin geliştirilmesi [Developing a scale for measurement and evaluation techniques]. *Hacettepe University Journal of Education,* (18), 42-50.

Baysen, E. (2006). Öğretmenlerin sınıfta sordukları sorular ile öğrencilerin bu sorulara verdikleri cevapların düzeyleri [The levels of teacher questions and student answers]. *Kastamonu Education Journal, 14*(1), 21-28.

Bektaş, M., & Kudubeş, A. A. (2014). Bir ölçme ve değerlendirme aracı olarak: Yazılı sınavlar [As a measurement and evaluation tool: Written exams]. *E-Journal of Dokuz Eylul University Nursing Faculty, 7*(4), 330-336.

Beydoğan, H. Ö. (1998). *Okullarda ölçme ve değerlendirme* [Measurement and evaluation in schools.]. Eser Ofset.

Black, P. (1995). Assessment and feedback in science education. *Studies in Educational Evaluation, 21*(3), 257-279.

Brookhart, S. M. (1997). Effects of the classroom assessment environment on mathematics and science achievement. *Journal of Educational Research, 90*(6), 323-331.

Büken, R., & Artvinli, E. (2020). Sosyal bilgiler ders kitaplarındaki coğrafya etkinliklerinin yenilenmiş Bloom taksonomisine göre analizi [Analysis of geography activities in social studies textbooks according to Bloom's revised taxonomy]. *The Journal of Turkish Educational Sciences, 19*(2), 1119-1142.

Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.), Pearson Education Inc.

Creswell, J., & Poth, C. (2018). *Qualitative inquiry and research design choosing among five approaches.* (4th ed.), SAGE Publications, Inc.

Çepni, S., & Ayvacı, H. Ş. (2007). Fen ve teknoloji eğitiminde alternatif (performans) değerlendirme yaklaşımları [Alternative (performance) evaluation approaches in science and technology education]. In S. Çepni (Ed.). *Kuramdan uygulamaya fen ve teknoloji öğretimi* [Teaching science and technology from theory to practice]. Pegem A.

Çepni, S., Özsevgeç, T., & Gökdere, M. (2003). Okullarda ve lise giriş sınavlarında sorulan fen bilgisi sorularının Bloom taksonomisine göre karşılaştırılması [Comparison of science questions asked in schools and high school entrance exams according to Bloom's taxonomy]. *Milli Eğitim Dergisi,* (157), 30-39.

Çolak, K. (2008). *Tarih dersi sınav sorularının Bloom taksonomisinin bilissel alan düzeyi açısından sınıflandırılması* [Classification of high school history examinations questions according to cognitive levels of Bloom taxonomy]. (Tez No. 231749). [Master Thesis, Karadeniz Teknik University], Natioanal Thesis Center.

Çopur, A., & Önal, H. (2023). Sosyal bilgiler dersi yazılı yoklama sorularının Bloom taksonomisi ve program kazanımlarına uygunluğu açılarından incelenmesi [Investigation of social studies course written attendance questions in terms of Bloom's taxonomy and suitability for program outcomes]. *E-International Journal of Educational Research*, *14*(4), 195-212.

Danışmaz, Ö. (2019). *Lise tarih dersi sınav sorularının Bloom taksonomisinin bilişsel alan basamaklarına göre sınıflandırılması* [The classification of high school history lesson exam questions according to cognitive domain of Bloom's Taxonomy]. (Tez No. 587052). [Master Thesis, Trabzon University], National Thesis Center.

Demirel, Ö. (2015). *Eğitimde program geliştirme kuramdan uygulamaya* [Curriculum development in education from theory to practice]. Pegem Akademi.

Downing, S. M. (2005). The effects of violating standard item writing principles on tests and students: The consequences of using flawed test items on achievement examinations in medical education. *Advances in Health Sciences Education,* 10, 133-143.

Durukan, E., & Demir, E. (2017). 6, 7 ve 8. Sınıf Türkçe dersi öğrenci çalışma kitaplarındaki etkinliklerin Bloom'un yenilenmiş taksonomisine göre sınıflandırılması [Classfying the activities in the Turkish lesson student workbook for the 6th, 7th and 8th grades students according to the renovated taxonomy of Bloom]. *International Journal of Turkish Literature Culture Education, 6*(3), 1619-1629.

Ertürk, S. (1975). *Eğitimde program geliştirme* [Program development in education]. Yelkentepe.

Gelbal, S., & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme konusunda yeterlilik algıları ve karşılaştıkları sorunlar [Teachers' proficiency perceptions of about the measurement and evaluation techniques and the problems they confront]. *Hacettepe University Journal of Education,* (33), 235-247.

Göçer, A. (2015). Lisansüstü eğitim gören Türkçe öğretmenlerinin yazılı sınav sorusu yazabilme yeterlikleri [Qualification for preparing written examination question of Turkish teachers who upper graduate education]. *Kafkas University Journal of the Institute of Social Sciences,* (17), 41-57.

Gömleksiz, M. N., & Bulut, İ. (2007). Yeni fen ve teknoloji dersi öğretim programının uygulamadaki etkililiğinin değerlendirilmesi [An assessment of the implementation of new science and technology curriculum]. *Hacettepe University Journal of Education*, *32*(32), 76-88.

Gömleksiz, M., & Erkan, S. (2014). Bilişsel alan davranışlarının ölçülmesi [Measurement of cognitive domain behaviors]. In B. Çetin, (Ed.), *Eğitimde ölçme ve değerlendirme* [Measurement and evaluation in education]. (pp. 67-142). Nobel Akademik.

Heverly, M., & Fitt, D. X. (1994). Classroom assessment of student competencies. *Assessment & Evaluation In Higher Education, 19*(3), 215-225.

Karamustafaoğlu, S., Sevim, S., Karamustafaoğlu, O., & Çepni, S. (2003). Analyses of Turkish high school chemistry examination questions according to Bloom's taxonomy. *Chemistry Education: Research and Practice, 4*(1), 25-30.

Kaya, E., Bal, D. A., & Sezek, F. (2002). Biyoloji eğitimini değerlendirmede kullanılan soru tipleri hakkında lise ve meslek lisesi öğrencilerinin görüşleri [The views of the high school and vocational high school students regarding the question types used to evaluate biology courses]. *Erzincan University Journal of Education Faculty, 4*(1), 55-64.

Kılıç, D. (2010). *Sosyal bilgiler öğretmenlerinin tarih konuları ile ilgili soru sorma becerilerinin Bloom taksonomisine göre değerlendirilmesi* [Evaluation of social studies teachers questioning skills about subject of history according to Bloom taxonomy]. (Tez No. 250818). [Master Thesis, Marmara University], National Thesis Center.

Köken, N. (2002). Sosyal bilgiler öğretiminde "soru sorma metodu"nun önemi [The importance of the "questioning method" in social studies teaching]. *Afyon Kocatepe University Journal of Social Sciences*, *4*(2), 91-104. https://sbd.aku.edu.tr/IV2/6-(91-104).pdf

Kutlu, Ö., Yalçın, S., & Pehlivanoğlu, E. B. (2010). İlköğretim programında yer alan kazanımlara dayalı soru yazma ve puanlama çalışması [A study on writing and scoring open-ended questions based on the primary school curriculum objectives]. *Elementary Education Online, 9*(3), 1201-1215.

Küçükahmet, L. (1999). *Öğretimde planlama ve değerlendirme* [Planning and evaluation in instruction]. Alkım.

Linn, R. L., & Gronlund, N. E. (1995). *Measurement and assessment in teaching.* (7th ed.). Prentice Hall.

McMillan, J. (2015). *Classroom assessment: principles and practice for effective standards-based instruction* [Sınıf içi değerlendirme etkili ölçütlere dayalı etkili bir öğretim için ilke ve uygulamalar]. (A. Arı, Trans.; 6th ed.). Eğitim Kitabevi.

Ministry of National Education [MoNE]. (2005). *Sosyal bilgiler dersi öğretim programı ve kılavuzu* [Social studies course curriculum and guide]. MEB.

Orbeyi, S., & Güven, B. (2008). Yeni İlköğretim Matematik Dersi Öğretim Programı'nın değerlendirme öğesine ilişkin öğretmen görüşleri [Teachers' perceptions related to the evaluation component of the education program of primary school education mathematics studies]. *Journal of Theory and Practice in Education*, *4*(1), 133-147. http://eku.comu.edu.tr/index/4/1/sorbeyi_bguven.pdf

Özden, Y. (1998). *Öğrenme ve öğretme süreçleri* [Learning and teaching processes]. Pegem Akademi.

Özgül, D. A., & Polat, K. (2021). Öğretmen adaylarının ortaokul matematik dersi kazanımlarını yenilenmiş Bloom taksonomisine yerleştirebilme ve soru yazabilme becerileri [Investigating of mathematic teacher candidates' skills to placement the learning outcomes according to revised bloom taxonomy and to write questions]. *Journal of Bayburt Education Faculty, 16(32)*, 304-328.

Patton, M. (2014). *Qualitative research & evaluation methods* [Nitel araştırma & değerlendirme yöntemleri]. (M. Bütün & S. B. Demir, Trans.; 3 rd ed.). Pegem Akademi.

Sağlam-Arslan, A., Avcı, A., & İyibil, Ü. (2008). Fizik öğretmen adaylarının alternatif ölçme değerlendirme yöntemlerini algılama düzeyleri [Physics teacher candidates' perception levels of alternative measurement and evaluation methods]. *Dicle University Journal of Education Ziya Gökalp Education, 11*, 115-128.

Safran, M. (2002). Ortaöğretim kurumlarında tarih öğretiminin yapı ve sorunlarına ilişkin bir araştırma [A research on the structure and problems of history teaching in secondary education institutions]. *Türk Yurdu Dergisi, 22*(175), 73-79.

Şanlı, C., & Pınar, A. (2017). Sosyal bilgiler dersi sınav sorularının yenilenen Bloom taksonomisine göre incelenmesi [An investigation of the social sciences courses exam questions according to revised Bloom's taxonomy]. *Elementary Education Online, 16*(3), 949-959.

Tarrant, M., Knierim, A., Hayes, S. K., & Ware, J. (2006). The frequency of item writing flaws in multiple choice questions used in high stakes nursing assessments. *Nurse Education Today,* 26, 662-671.

Tekin, H. (1994). *Eğitimde ölçme ve değerlendirme* [Measurement and evaluation in education]. Yargı.

Tekin, H. (2004). *Eğitimde ölçme ve değerlendirme* [Measurements and evaluation in education]. Yargı.

Uymaz, M., & Çalışkan, H. (2019). Öğretmen yapımı sosyal bilgiler dersi sınav sorularının soru türleri ve kapsam geçerliğine göre incelenmesi [An investigation on the teacher-made social studies course exam questions in terms of the type of the questions and content validity]. *International Journal of Humanities and Education, 5*(10), 77-113.

Yavuz, Ö. Ç., & Furkan, E. (2020). Sınav hazırlama ve yönetimi üzerine bir uygulama [An application on question preparation and management]. *Journal of Information Systems and Management Research, 2*(2), 10-15.

Yıldırım, A., & Şimşek, H. (2011). *Sosyal bilimlerde nitel araştırma yöntemleri* [Qualitative research methods in the social sciences] (8th ed.). Seçkin.

Yıldız, S. B. (2004). *Öğretmenler için soru sorma sanatı* [The art of asking questions for teachers]. Asil.

Yıldız, İ., & Uyanık, N. (2004). Matematik eğitiminde ölçme değerlendirme üzerine [On measurement and evaluation in mathematics teaching]. *Gazi Üniversitesi Kastamonu Eğitim Dergisi, 12*(1), 97- 104.

Yılmaz, A. A., & Gazel, A. (2017). 4. ve 7. Sınıf sosyal bilgiler derslerinde sorulan öğretmen sorularının Bloom taksonomisinin bilişsel alanına göre incelenmesi [The study of teacher's questions asked in 4th and 7th social sciences classes according to Bloom taxonomy's cognitive domain]. *Afyon Kocatepe University Journal of Social Sciences, 19*(2), 173-186.

Yılmaz, A., Akbaba, F. D., Halıpınar, F. M., Oral, S., & Ulusoy Ünlü, A. (2021). Okul öncesi eğitim etkinlik kitabının yenilenmiş Bloom taksonomisine göre incelenmesi [Review of a pre-school activity book according to Bloom's revised taxonomy]. *Journal of Uludag University Faculty of Education, 34*(1), 343-385.

BIOGRAPHICAL NOTES

Contribution Rate of Researchers

Author 1: 50%

Author 2: 50%

Conflict Statement

As the authors of the study, we state that we do not have any conflict of interest.

Genişletilmiş Türkçe Özet

Sosyal Bilgiler Dersi Yazılı Yoklama Sorularının Soru Türü ve Hazırlama Teknikleri Açısından İncelenmesi

Giriş

Öğretmenler, öğrenme çıktılarını doğru değerlendirebilmek için, doğru verileri elde etmelidir. Doğru veri ise doğru ölçme araçları ile toplanabilir. Dolayısıyla, her yapılan sınav, doğru bir değerlendirme yapıldığı anlamına gelmeyebilmektedir. Bu sorunun nedenleri olarak, soru düzeyi ve kapsam geçerliliği gibi doğrudan içerikle ilgili sorunlar olabileceği gibi soruyu hazırlayandan kaynaklanan, hatalı sorma da gösterilebilir. Öğretmenler, sınav hazırlarken genellikle çoktan seçmeli, açık uçlu, iki seçili, eşleştirme, kısa cevaplı ve boşluk doldurmalı gibi soru tiplerini kullanmaktadır. Sınav evrakında sadece bunlardan biri ile yapılmış bir ölçme aracı hazırlanabileceği gibi, birden fazla soru tipi de bir arada kullanılabilmektedir. Göçer'e (2015) göre hazırlanan sınavın, gerçek amacına ulaşabilmesi için, içerik açısından olduğu kadar, biçimsel açıdan da doğru yazılması gerekir. Bu nedenle öğretmenlerin soru hazırlayabilme becerilerine sahip olması değerlendirmenin sağlıklı yapılabileceği bir ölçme aracı oluşturmanın önemli bir aşamasıdır.

Kuralları dikkate almadan ve acele olarak hazırlanan sorular, içerik ve kapsam kadar, soru yazımında birçok yazım hatası da içerebilmektedir (Yavuz & Furkan, 2020). Ölçme için hazırlanan sınavdaki sorulara ait maddelerin her biri, ölçülmesi beklenen becerileri tam olarak ölçebilmelidir (Küçükahmet, 1999). Bunu için de yazılan her maddenin, amaca ve kurallara uygun olması gerekir. 2005 yılından itibaren uygulanmaya başlanan yapılandırmacı yaklaşımların ön plana çıkarılması süreci ile, ölçme araçları da çeşitlenmeye başlamıştır. Öğretmenlerin, bu yeni ölçme araçlarının varlığından haberdar oldukları kadar, bunların hazırlanış ve uygulanış süreçlerini de bilmeleri gerekmektedir. Zira değerlendirmenin sağlıklı olması, ölçme aracının sağlıklı kullanımına bağlıdır. Çepni vd., (2003), farklı soru türlerinin kullanılmasının faydalarından bahsederken, bunun öğrencilerin düşünme düzeylerini de etkilediğini vurgulamışlardır. Sorular hem öğrencilerin düşünmelerini etkilediği gibi hem de derse karşı

olan tutumlarını ve akademik başarılarını da ölçtüğünden her zaman araştırmacıların ilgisini çekmişlerdir (Akpınar & Ergin, 2006). Linn ve Gronlund (1995), soru hazırlamayı, belirli bilgileri ve becerileri gerektiren bir sanat olarak tanımlamışlar ve eğitimin tüm süreçlerinde olduğu gibi soru hazırlamanın da eğitiminin alınması gerektiğini vurgulamışlardır. Bu hususta (Karamustafaoğlu vd., 2003) eğitimin kalitesinin artırılmasında öğretmenin rolüne değinirken, kaliteli sorular hazırlayabilen öğretmenlerin bu süreçte daha fazla katkı sağladıklarını vurgulamışlardır. Tekin (2004), öğretmenlerin soru hazırlama sırasında, ilgili kazanımlar için ne tür soru çeşidi seçileceğine göre değil, genellikle en iyi bildikleri soru türlerini tercih ettiklerini belirtmiştir. Aslında birçok araştırmada da bu durum, farklı yorumlanarak, bunların birer hata olduğu ve öğretmenlerin soru hazırlama ve soru tipi belirleme konularında hizmet içi eğitimlere alınması gerektiği şeklinde yorumlanmaktadır (Anıl & Acar, 2008; Bakırcı vd., 2022;Gelbal & Kelecioğlu, 2007).

Türkiye'de öğretmenlerin hazırlayıp uyguladıkları soruların incelendiği çalışmalara bakıldığında, neredeyse tamamına yakınının, soruların kazanımlarla ilişkilendirilmesi (Aydın ve Uçgun, 2020; Kutlu vd., 2010; Özgül ve Polat, 2021) veya düzeylerinin belirlenmesi (Akpınar ve Ergin, 2006; Baysen, 2006; Büken ve Artvinli, 2020; Çepni vd., 2003; Karamustafaoğlu vd., 2003; Özgül ve Polat, 2021) üzerine oldukları görülmektedir. Bu çalışmada ise Sosyal Bilgiler öğretmenlerinin kendi hazırladıkları sınavlardaki soru tiplerinin nelerden oluştuğu ile, kullanılan soru tiplerinin, soru ve madde yazım kuralları açısından değerlendirilmesi yapılmaya çalışılmıştır. Bu doğrultuda çalışmada; yazılı yoklama sorularının soru tiplerine göre dağılımı ve soru yazım tekniklerine uygunluğu nasıldır? Sorularına yanıt aranacaktır.

Yöntem

Çalışmada nitel araştırma yöntemlerinden birisi olan doküman analizi kullanılmıştır. Toplanan veriler, içerik analizine de tabi tutulmuştur. Çalışmada doküman incelemesi süreci, uygun dokümanları bulma, orijinallik kontrolü, kod-katalog oluşturma sistematiği belirleme ve verilerin analizi (içerik analizi) aşamaları (Merriam, 2009) referans alınarak yürütülmüştür. Araştırmanın problemi temel alınarak doküman incelemesi tekniğinin probleme ışık tutacak verilerin toplanması için yeterli olacağına karar verilmiştir. Araştırma grubu Balıkesir ili Altıeylül ve Karesi ilçelerinde görev yapan sosyal bilgiler öğretmenleri arasından amaçlı örnekleme tekniklerinden uygun durum örneklemesi ile belirlenmiştir. Amaçlı örnekleme tekniği, çalışmaya gönüllü ve uygun katılımcıların tespitine olanak sunduğundan (Patton, 2014) tercih edilmiştir. Bu doğrultuda çalışma, 20 gönüllü sosyal bilgiler öğretmeniyle yürütülmüştür.

Çalışmada veriler, araştırmacı tarafından literatürden (McMillan, 2015) yararlanılarak geliştirilen, "Soru Yazım İlkeleri Kontrol Çizelgesi" (Ek-1) aracılığıyla incelenmiştir. Taslak olarak hazırlanan çizelge dil ve içerik açılarından uzman görüşüne sunulmuştur. Bu kapsamda dil kontrolü için Türk Dili ve Edebiyatı (1) uzmanından, içerik için ise ölçme ve değerlendirme (1) ve sosyal bilgiler (1) uzmanından görüş alınmıştır. Nihayetinde çoktan seçmeli, açık uçlu, iki seçimli, eşleştirme, kısa cevaplı ve tamamlama (boşluk doldurma) tipi sorular için 45 maddeden oluşan bir soru yazım kontrol aracı geliştirilmiştir. Araştırma kapsamında elde edilen veriler araştırmacı tarafından, farklı zamanlarda, iki defa analiz edildikten sonra verilerin bir bölümü (%30) bir alan uzmanı tarafından da analiz edilmiştir. Son olarak araştırmacı ve alan uzmanın analiz sonuçları mukayeseli şekilde değerlendirilerek, mutabık kalınan ve ayrı düşülen noktalar belirlenmiştir. Bu aşamadan sonra ortak bir analiz tablosu oluşturularak işlem sonlandırılmıştır.

Çalışmanın güvenirliğini sağlamak amacıyla verilerin, farklı zamanlarda, tekrar kontrolü yapılmıştır. Ek olarak, analiz edilen verilerin çapraz kontrolü, bir alan uzmanı tarafından daha yapılmıştır. Alan uzmanından, verilerin %30'unu analiz etmesi istenmiştir. Bu iki analizin analiz sonucundaki uyum (Miles ve Huberman, 2016) %92 olarak hesaplanmıştır.

Bulgular

Araştırma bulguları incelendiğinde yazılı yoklama sınavlarında öğretmenlerin her sınıf düzeyinde de en fazla çoktan seçmeli soru tipini tercih ettiği görülmektedir. Bu durum öğrencilerin çoktan seçmeli soruların yer aldığı genel sınavlara girecek olması ve bu tip soruların objektif ve hızlı değerlendirmeye imkân sunmasıyla ilgili olabilir. Çoktan seçmeli soru tipini sırasıyla iki seçkili, boşluk doldurma ve eşleştirme izlerken en az açık uçlu soru tipi tercih edilmektedir. Beşinci sınıf düzeyinde iki ve altıncı sınıf düzeyinde dört olmak üzere toplam altı tane açık uçlu soru kullanılması açık uçlu soruların değerlendirilmesinin zor olması ve öğrencilerin bu tür soruları cevaplamakta zorlanması gibi gerekçelerden kaynaklanabilir.

Araştırmadan elde edilen bir başka bulguya göre çoktan seçmeli soruların; "Soru kökü mümkün olduğunca açık, doğrudan ve basit bir şekilde ifade ediliyor mu? "Soru kökünde olumsuz ifadelerin kullanımından kaçınılmış mı?" "Dilbilgisi yapısı, uzunluk ve karmaşıklık konusunda tüm seçenekler aynı doğrultuda mı?" ve "Doğru cevabı bilmeyen öğrenciler için çeldiriciler mantıklı mı?" kriterleri açısından geliştirilmeye açık olduğu söylenebilir. Ayrıca "Seçenekler kısa mı? "Karmaşık seçeneklerden kaçınılıyor mu? ve "Doğru cevaplar, tüm seçenekler içinde eşit olarak dağıtılıyor mu?" kriterleri açısından soruların yazım tekniklerine önemli ölçüde uygunluk gösterdiği görülse de kriterlere uymayan sorular da tespit edilmiştir.

Araştırmadan elde edilen bulguların açık uçlu soruların; "Hedeflenen muhakeme becerileri ölçüldü mü? ve "Görev açıkça belirtildi mi?" kriterleri açısından geliştirilmeye açık olduğuna işaret ettiği ifade edilebilir. Ayrıca "Soruları cevaplamak için verilen süre yeterli mi? ve "Çeşitli sorular arasından seçim yapma durumunun önüne geçildi mi?" kriterleri açısından soruların yazım tekniklerine önemli ölçüde uygunluk gösterdiği görülmekle birlikte bu soru tipinde çok az sayıda soruya yer verilmesi daha kapsamlı bulgulara ulaşmanın önünde engel teşkil ettiği söylenebilir.

Elde edilen bir başka bulguya göre iki seçimli soruların; "Madde tek bir önerme ya da fikir içeriyor mu? "Doğru cevapların bir modeli var mı?" ve "Madde öğrencileri kandırmaya çalışmayacak şekilde yazılmış mı?" kriterleri açısından geliştirilmeye açık olduğu söylenebilir. Ayrıca "Madde Olumlu olarak ifade ediliyor mu?", "Maddedeki her iki ifadenin uzunluğu yaklaşık olarak aynı mı?" ve "Anlaşılır terimler kullanılıyor mu?" kriterleri açısından soruların genel olarak yazım tekniklerine uygunluk gösterdiği ifade edilebilir. Bununla birlikte örnek 7 ve 8 de görüleceği üzere ilgili kriterlere uymayan sorular da yer almaktadır.

Araştırmadan ulaşılan bulgulardan hareketle eşleştirme tipi soruların; "Her bir karşılığın bir ya da birden fazla sefer veya hiç kullanılmayabileceği açık mı?" ve "Öncüllerden daha fazla karşılık var mıdır?" kriterleri açısından geliştirilmeye açık olduğu ifade edilebilir. Ayrıca "Öğrencilerin cevapları nereye ve nasıl yerleştireceği açık mı?", "Her bir öncül için bir tane olası cevap var mıdır?" ve "Öncüller ve karşılıkları aynı sayfada mıdır?" kriterleri açısından soruların yazım tekniklerine büyük oranda uygunluk gösterdiği ifade edilebilir.

Elde dilen bulgular kısa cevaplı soruların; "Soru tek bir doğru cevabı olacak şekilde yapılandırıldı mı?" ve "Soru mümkün olduğunca kısa ve öz yazıldı mı?" kriterleri açısından geliştirilmeye açık olduğunu işaret etmektedir. Ayrıca "Kitaptan doğrudan alıntı soru sorulmasından kaçınıldı mı?" ve "Soruda kullanılan sözcükler öğrenciler için anlaşılır zorlukta mı?" kriterleri açısından soruların genel olarak yazım tekniklerine uygunluk gösterdiği söylense de "Soru mümkün olduğunca kısa ve öz yazıldı mı?" maddesinin özellikle 5. sınıflar için sorunlu olduğu da görülmektedir.

Araştırma bulguları incelendiğinde boşluk doldurma (tamamlama) sorularının; "Öğretim materyallerinden motamot dil kullanımı engellendi mi?", "Bırakılan boşluk cümlenin sonunda mı?", "Boşlukların hepsi aynı genişlikte mi?" ve "Doğru cevaba dair ipucu vermekten sakınıldı mı?" kriterleri açısından geliştirilmeye açık olduğu görülmektedir. Ayrıca "Bilgi değerlendirildi mi?" ve "Kısa ve öz cevap istendi mi?" kriterlerinin tümü açısından soruların yazım tekniklerine pek uygunluk gösterdiği söylenemez.

Tartışma ve Sonuç

Araştırma bulguları öğretmenlerin en çok tercih ettikleri soru tipinin çoktan seçmeli sorular olduğunu ortaya koymaktadır. Çoktan seçmeli soru tipini sırasıyla iki seçkili, boşluk doldurma ve eşleştirme izlerken en az açık uçlu soru tipi tercih edilmektedir. Araştırma bulgularıyla benzer sonuçlara ulaşan Şanlı ve Pınar (2017) sosyal bilgiler dersi sınav sorularını inceleyen çalışmalarında çoktan seçmeli soruları öğretmenlerin en çok tercih ettiği soru tipi olurken açık uçlu sorular en aza tercih ettikleri soru tiplerin biri olduğunu tespit etmişlerdir. Öğretmenlerin yazılı yoklama sınavlarında açık uçlu sorulara çok az sayıda yer vermesi dikkate değer görülmektedir. İncelenen az sayıda açık uçlu sorudan yola çıkarak bir yargıya varmanın zorluğuna karşın bu soruların hedeflenen muhakeme becerilerini ölçmekten uzak olduğu ve sorularda görevin açıkça belirtilmediği görülmüştür. Bu bulgular, sosyal bilgiler (Kılıç, 2010; Şanlı ve Pınar, 2017; Uymaz ve Çalışkan, 2019), tarih (Danışmaz, 2019) ve Türkçe (Aydın ve Uçgun, 2020) öğretmenlerinin en fazla tercih ettiği soru türünün çoktan seçmeli sorular olduğu bulgularıyla örtüşürken; öğretmenlerin sınavlarda çoğunlukla kısa cevaplı (Safran, 2002) ve açık uçlu (Çolak, 2008) sorular sorduğu sonucuna ulaşan çalışmalarla çeliştiği söylenebilir. Ayrıca merkezî seçme ve yerleştirme sınavlarında bu soru türünün kullanılması öğretmenlerin çoktan seçmeli soru türünü daha çok tercih etmelerinin bir diğer gerekçesi olarak düşünülebilir. Açık uçlu soruların ise az tercih edilmesi; öğrencilerin bu tür soruları cevaplamakta zorluk çekmesi ve açık uçlu soruların değerlendirilme zorluğundan (McMillan, 2015) kaynaklanabilir. Ek olarak, öğretmenlerin soru tipini belirlerken ilgili kazanımı ölçmeye en uygun soru tipini değil, genellikle en iyi bildikleri soru türlerini tercih etmeleri (Tekin, 2004) bu durumun bir başka gerekçesi olabilir.

Elde edilen bulgular çoktan seçmeli soruların yazımında; soru kökünü yalın olarak ifade etme, olumlu olarak yazma, dilbilgisi ve uzunluk bakımından eş değer madde oluşturma ve çeldirici madde yazımı açılarından öğretmenlerin geliştirilmeye ihtiyaç duyduğuna işaret etmektedir. Bulguları destekler nitelikte Downing (2005) dört testteki 219 maddenin incelenmesi üzerine yaptığı araştırmasında testlerin % 35 ile % 65 oranında madde yazım ilkesini ihlal ettiği ve iki testte en çok görülen ihlalin maddenin olumsuz yazılması olduğunu belirlemiştir. Tarrant vd.'nin (2006) doğru yanıtın çeldiricilerden daha uzun olması en sık rastlanılan ihlaller arasında yer aldığını belirlemişlerdir. Bu buldular da çalışma bulgularıyla örtüşmektedir. Aybek, Kartal ve Yıldırım (2022) ise fen bilgisi ve sosyal bilgiler dersinde sorularını madde yazım ilkelerine uygunluğu açısından incelendikleri çalışmalarında hem fen bilgisi hem de sosyal bilgiler alanında maddelerin büyük bir çoğunluğunda ana

fikrin madde kökünde açık ve anlaşılır bir biçimde ifade edildiğini tespit etmişleridir. Bu tespitlerin mevcut araştırmanın soru kökünü yalın olarak ifade etme ile ilgili bulgularıyla kısman örtüştüğü söylenebilir. Ek olarak iki seçkili soruların yazımında öğretmenlerin; maddeleri tek bir önerme veya fikir etrafında şekillendirme ve doğru cevapları bir modele göre tasarlamama gibi ölçütlerde geliştirilmeye ihtiyaç duymaktadırlar. Benzer bulgulara ulaşan Aybek vd. (2022) çalışmalarında birleşik yanıt vermeyi gerektiren maddelere yer verildiği tespit etmişlerdir. Eşleştirme tipi soruların yazımında ise; her bir karşılığın bir ya da birden fazla sefer veya hiç kullanılmayabileceğini belirtme ve öncüllerden daha fazla karşılığa yer verme kriterleri öğretmenlerin geliştirilmeye açık olduğu diğer yönler olarak ön plana çıkmıştır. Ayrıca tamamlama (boşluk doldurma) sorularında; öğretim materyallerinden motamot dil kullanımı, boşlukların konumlandırılması ve yapılandırılması ile doğru cevaba dair ipucu verme gibi ölçütlerde sorunlar olduğunu ortaya koymaktadır.

Ölçme için hazırlanan sınavdaki sorulara ait maddelerin her biri, ölçülmesi beklenen becerileri tam olarak ölçebilmelidir (Küçükahmet, 1999). Bunun için de yazılan her maddenin içerik açısından olduğu kadar, biçimsel açıdan da doğru yazılması gerekir (Göçer, 2015). Kuralları dikkate almadan ve acele olarak hazırlanan sorular, içerik ve kapsam kadar, soru yazım tekniği açısında birçok yazım hatası da içerebilmektedir (Yavuz & Furkan, 2020). Aslında birçok araştırmada da bu durum, farklı yorumlanarak, bunların birer hata olduğu ve öğretmenlerin soru hazırlama ve soru tipi belirleme konularında hizmet içi eğitimlere alınması gerektiği şeklinde yorumlanmaktadır (Anıl & Acar, 2008; Bakırcı, Gelbal & Kelecioğlu, 2007; Özmantar, & Bozkurt, 2022). Ek olarak farklı soru türlerinin kullanılması öğrencilerin düşünme düzeylerini (Çepni, Özsevgeç ve Gökdere, 2003) etkilediğinden hareketle, yazılı yoklamaların, sınav soru yazım tekniklerine uygun hazırlanmasının ayrı bir öneme sahip olduğu söylenebilir. Bu hususu destekler nitelikte bulgulara ulaşan Karamustafaoğlu vd. (2003), eğitimin kalitesinin artırılmasında öğretmenin rolüne değinirken, kaliteli sorular hazırlayabilen öğretmenlerin bu süreçte daha fazla katkı sağladıklarını vurgulamışlardır. Bu konuda Balcı ve Tekkaya (2000) ise öğretmen yapımı sınavların öğretmenlerin sınav hazırlama süreçleri ve kurallarına uymamaların nedeniyle bu sınavların değerlendirmesinin çok da sağlıklı olmadığını belirtmişlerdir. Buraya kadar zikredilen çalışmalardan elde edilen bulguların, soru hazırlama tekniklerine uygun soru hazırlamanın önemine işret etme bakımından, mevcut araştırma bulgularını destekler nitelikte olduğu söylenebilir.

Öneriler

Araştırmadan elde edilen bulgulardan hareketle;

1. Öğretmenlerin soru türleri ve yazım ilkeleri konularında yeterliklerini artırıcı eğitimler yapılması,
2. Öğrencilerin üst düzey düşünme becerilerinin geliştirilmesi ve analiz, sentez ve değerlendirme basamağında olan kazanımlara ulaşma düzeylerinin belirlenmesi açılarından son derece önemli olan açık uçlu soruların çok az kullanılmasının nedenleri araştırılarak çözümüne yönelik uygulamalı çalışmalar yapılması,

Ölçme için hazırlanan sınavdaki sorulara ait maddelerin her biri, ölçülmesi beklenen becerileri tam olarak ölçebilmelidir (Küçükahmet, 1999). Bunun için de yazılan her maddenin içerik açısından olduğu kadar, biçimsel açıdan da doğru yazılması gerekir (Göçer, 2015). Bu sonuçlardan yola çıkarak öğretmenlerin soru hazırlama teknikleri açısından mesleki gelişimini sağlamak üzere projeler yapılması yönünde öneriler getirilmiştir.