e-Kafkas Journal of Educational Research e-KJER Research Article

Investigation of the Change in the Features of Turkish Items in the High School Entrance System 2018-2022

Ayfer Sayın[1] **Deniz Melanlıoğlu**[2]

To cite this article:
Sayın, A. & Melanlıoğlu, D. (2023). Investigation of the change in the features of Turkish items in the High School Entrance System 2018-2022. *e-Kafkas Journal of Educational Research, 10*, 652-671. doi: 10.30900/kafkasegt.1377107

Research article **Received:** 16.10.2023 **Accepted:** 29.12.2023

Abstract

This study aims to examine the structural features of Turkish items in the High School Entrance System (LGS), first implemented in 2018, and to determine changes in these features over the years. This descriptive survey initially investigates how 100 Turkish items administered in the LGS from 2018 to 2022 changed in terms of test difficulty, discrimination, and internal consistency. Subsequently, textual features (sentence, word, and syllable length; readability), contextual features (visuality, meaning of the item stem, common stem), and purpose features (content domain, text type) of the items were determined, and changes in these features over the years were examined. As a result of the study, it was found that the test difficulty of the LGS Turkish test decreased over the years, indicating an increase in item difficulty. While there was no significant overall change in test discrimination, a general decrease in test discrimination was observed. The study also revealed variations in item lengths over the years, with an observed increase in the overall test's average length during the same period. The content of the items predominantly consisted of informative texts; however, in some years, poetry and narrative texts were included in the items.

Keywords: Turkish education, High School Entrance System, reading comprehension items, structural features, readability.

[1] Ayfer Sayın, ayfersayin@gazi.edu.tr, Gazi University

[2] Deniz Melanlıoğlu, Istanbul University

Introduction

The Turkish language course, in which mother tongue education is carried out systematically, aims to provide primary and secondary school language skills. Students' competence in language skills is directly proportional to their academic success (Sevgi & Karakaya, 2021; Şahin, 2022). The performance indicators for the courses are the result of measurement and evaluation activities. These indicators are also one of the main variables determining students' transition between levels. In Turkey, two central exams determine the transition between levels. The first one is the High School Entrance System (LGS), which regulates the transition from secondary school to high school, and the second one is the Turkish Proficiency Test (known as TYT) and the Field Proficiency Test (known as AYT) used for university entrance. It can be said that LGS has particular importance in terms of the average age of the students and that it is the first centralized exam they face. Students enrolled in the eighth grade of formal schools participate in this verbal and numerical exam. With the exam results, students can be placed in science high schools, social sciences high schools, project schools, and Anatolian technical programs of vocational and technical Anatolian high schools (MEB-LGS Report, 2022, p. 12). The number of Turkish items and the coefficient for scoring are higher in the verbal section than in other verbal subtests. The Turkish items in the central exam are based on the eighth-grade outcomes of the Turkish Curriculum (for grades 1-8) (MEB-LGS Report, 2022, p. 13; MEB-LGS Report, 2021). Although the curriculum includes learning outcomes for listening/monitoring, speaking, reading, and writing skills, the central exams focus on measuring specific learning outcomes (Aydın, 2022; Calp & Alpkaya, 2021; Taşıyaran, 2022).

In line with the Turkish Lesson Teaching Program (for Grades 1-8), the education provided during secondary school serves the purpose of using language skills effectively, making reading and writing skills a habit, developing high-level thinking skills, using information sources effectively, reinforcing national and spiritual values, and gaining aesthetic pleasure (MEB, 2019a). The fact that reading and writing skills are given particular importance in the program can be explained by the fact that both are acquired during the school process. For this reason, it is possible to say that what is expected from students in the central exams is to realize the reading and partial writing achievements. Of course, it is known that this situation is affected by many variables, such as the multiple-choice preparation of central exams, the breadth of the target group taking the exam, and the fact that reading is considered the essential element of personal learning. According to Uluyor and Eryılmaz (2015), among these variables, reading is more effective in raising individuals who can think creatively and critically, cooperate, communicate effectively, take responsibility, try different ways of accessing information, and utilize technology for its purpose. For students, reading is seen as a great source of input in acquiring world knowledge (Januarty & Nima, 2018). According to Hiebert (1983), reading is an essential life skill, and it is not possible to be successful in life without the ability to read well. Accordingly, measuring students' reading skills in central transition exams is essential.

Reading occurs through an interaction between author-reader-text. Many factors, such as prior knowledge, cultural background, value judgments, and expectations, play an active role in realizing this communication (Wigfield & Guthrie, 1997). In this regard, meaning-making in reading is a process in which the reader puts his/her prior knowledge and mental skills to work; variables such as the reader's prior knowledge, interest, need, and reading level can affect the meaning-making process positively or negatively (Melanlıoğlu, 2021). Guthrie and Scafiddi (2004) discuss the factors affecting the meaning-making process of reading as text, reader, the interaction between the reader and the text, and the mental state of the reader after reading the text; Singer (1978) examines them under four main headings: individual differences, reading task, reading purpose and text. Uyar (2015) draws attention to the reader (the reader's prior knowledge and the reading strategy used), motivation, text, and context elements that affect the process. These elements can be considered as determinants of the level of reading comprehension. Reading comprehension refers to the comprehension of the reading material (Özdemir, 2011). The stages that need to be realized to reach comprehension of the reading material are as follows (Davis, 1944):

- Vocabulary
- Predicting the meaning of words or phrases based on context
- Follow the organization of the text and identify antecedents and references within it
- Select the main idea of the text
- Be able to answer items specifically answered in the text
- To be able to answer items that are answered in the text but not in the words in which the item is asked
- Making inferences from the text
- To be able to recognize literary arts in the text
- Identifying an author's purpose, intention, and point of view, making inferences about an author

Considering the nine stages listed, it should only be expected to follow them in some reading texts. For example, when an informative text is read, there is no need to realize the stage of recognizing literary arts to establish meaning in this text. Therefore, the steps followed in reading comprehension differ according to the text type. However, vocabulary knowledge, the first stage, is necessary for making meaning in every text type, including informative, narrative, and poetry. Orasanu (1986) asserts that vocabulary knowledge is the principal determinant influencing the information a reader brings to the text, the comprehension of the text, and the subsequent capacity for learning and retention. There needs to be more than the reader's vocabulary knowledge to understand the text; the reader should also have fluent reading skills. Reading fluency is associated with cognitive processes defined as information retrieval, comprehension, evaluation, and reflection (OECD, 2019), and it includes the speed and accuracy of reproducing text into spoken language (Jiang, Sawaki & Sabatini, 2012). Fluency improves as readers progress in automatic decoding and word recognition, resulting in more accurate and faster reading. As readers develop automaticity in word recognition, text comprehension improves at the cognitive level (Perfetti & Hogaboam, 1975; Stanovich, 1980). For reading fluency to be realized, readers must have automatic recognition, vocabulary knowledge, formal discourse structure knowledge, world knowledge, and the ability to use cognitive and metacognitive strategies (Grabe, 1991). Gabb (2000) asks an essential item about why students face obstacles in moving to the fluency stage despite having basic decoding skills. At this point, variables such as the type of text encountered, readability level, and vocabulary knowledge should not be ignored. Reading fluency is also considered essential to meaning-making in PIRLS and PISA, which are organized internationally and focus on reading comprehension.

In the PISA application, in which students in the same age group as LGS students participate, six competency areas for reading literacy are defined. These competency areas, which are shaped according to the text and the behavior expected from the student, follow a hierarchy. The items in the LGS Turkish test are similar to the reading items in the PISA application. It is known that the skills that the items in the LGS test aim to focus on a higher level than the previous exam - TEOG- items (Ayyıldız & Aktaş, 2022; Azili & Tutkun, 2021; Çepni, 2019; Kızkapan & Nacaroğlu, 2019; Şan & İlhan, 2022). In the 2023 Education Vision Document published by the Ministry of National Education, it is stated that "New generation digital measurement materials that support metacognitive skills will be developed so that students can achieve the desired results in international exams such as PISA." (MEB, 2018a). In this context, it is seen that Turkish items were prepared per visual reading and context-based item logic.

In this context, LGS Turkish items focus on students' analytical and critical thinking, interpretation, and evaluation skills. Batur, Ulutaş, and Beyret (2019) and Çiçek and Dilekçi (2022) found that LGS Turkish items met the PISA reading skills objectives at levels 2, 3, and 4, and explained the reason for this result with the fact that LGS was not prepared at the targeted level in terms of measuring high-level thinking skills. This finding also coincides with the results of Aktaş's (2022) study in which LGS items were associated with PISA reading proficiency levels. Kanık Uysal (2022) determined that there were no items at the 5th and 6th levels, characterized as high-level reading

skills according to PISA reading skills proficiency levels, in LGS. In addition, the researcher, who also evaluated the LGS items according to the revised Bloom's taxonomy, stated that there were only items that could be associated with the creation step for higher-order thinking skills and that the analysis and evaluation steps were not considered. In this sense, Vural (2020) states that the existing exam cannot measure higher-order thinking skills. Gökdemir, Aydaşgil, and Topçuoğlu (2021) also found that the items in the 2020 LGS mainly corresponded to the comprehension and recall levels of the taxonomy. Erden (2020), who reveals the effect of the "new generation items" defined at the point of measuring higher-order thinking skills of the exam in the field in line with teacher opinions, states that Turkish teachers find the new generation items focused on higher-order thinking skills suitable for reading and grammar achievements, they like them in terms of associating them with higher-level achievements, but they find them insufficient in terms of the items being above the achievements, focusing on specific achievements, not including verbal logic items in the achievements, trying to make Turkish items similar to mathematics items and not addressing all learning areas. This result is similar to the results of Aydın (2022), who found that the LGS Turkish items conducted between 2018 and 2021 did not include listening and speaking objectives at all and measured 85% reading and 15% writing objectives. Based on the fact that LGS Turkish items were prepared by taking 8th-grade achievements into consideration, Calp and Alpkaya (2021) concluded in their study that LGS Turkish items were sufficient to measure the level of achievement of the learning outcomes in the Turkish curriculum, but ignored to measure the achievements related to some skill areas.

Studies on LGS Turkish items (Aktaş, 2022; Altun, 2021; Batur, Ulutaş & Beyrut, 2019; Benzer, 2019; Calp & Alpkaya, 2021; Çiçek & Dilekçi 2022; Ekinci & Bal, 2019; Erden, 2020; Kanık Uysal, 2022; Kılkapan & Nacaroğlu, 2019; Ordu, Engin & Topçuoğlu, 2021; Sayın & Takıl, 2023; Soysal & Güngör, 2022;), it is seen that LGS Turkish items are addressed from various perspectives such as reflecting the learning outcomes in the Turkish curriculum, representing language skills, testing higher-order thinking skills, and meeting the reading levels in international exams. The correct answer rate of the items and the variables affecting this rate (such as the educational level of parents) are discussed in the reports published by the Ministry of National Education on the evaluation of LGS (MEB-LGS Report, 2022; MEB-LGS Report, 2021). When the literature on the subject was reviewed, no correlational studies were found on LGS Turkish items from 2018 to date.

Research Questions

The current study aims to examine the structural and psychometric features of LGS Turkish items between 2018 and 2022 in detail and to determine the relationship between these two factors. In line with the stated purpose, the questions sought to be answered within the scope of the research can be listed as follows:

1. How do LGS 2018-2022 Turkish test statistics (test difficulty, discrimination, internal consistency) change?
2. How do the textual features (sentence, word, syllable length, and readability) of LGS 2018-2022 Turkish test items change?
3. How do the contextual features of LGS 2018-2022 Turkish test items (visuality, meaning of the item root, common stem) change?
4. How do the purpose features (sub-topic, text type) of LGS 2018-2022 Turkish test items change?
5. How do the test statistics, textual, contextual, and purpose features of LGS 2018- 2022 Turkish test items change together?

Considering the listed sub-problems, it is aimed to examine the change in test statistics in the context of test difficulty index, test discrimination, and internal consistency of the items; to reveal the structural features by considering textual, content, and objective qualities; and to determine the change in these features. Considering that 1,031,799 (83.46%) of the 1,236,308 eighth-grade students who automatically applied to the 2022 LGS (MEB, LGS-2022 Report, p. 18) participated in the exam (MEB, LGS-2022 Report, p. 18), it is believed that the results of the research are essential for policymakers, teachers, students, parents, and interested parties.

Method

Research Design

Since this study aims to reveal the structural features of LGS Turkish items and determine how these features change over the years, the research was conducted using a descriptive survey research method. Descriptive survey research aims to summarize the features of a given situation (LGS Turkish items) by describing its features as thoroughly and carefully as possible and revealing its features in detail (Büyüköztürk, et al., 2020).

Sample/Documents

In the study, MoNE compiled the changes in LGS Turkish test statistics from post-exam reports. The textual (sentence, word, and syllable length; readability), contextual (visuality, meaning of the item stem, common stem), and purpose features (sub-topic, text type) of the items were analyzed by the researchers. All documents were accessed from MoNE's official website (https://raporlar.meb.gov.tr/); no participants were involved. All human research and all other activities that involve partially human research, regardless of whether there is a sponsor/supporter, must be reviewed and approved by the relevant ethics committee before data collection can begin. All research involving human participants (including interviews, questionnaires, and questionnaires) must be assessed by the relevant ethics committee, except for research that does not require ethics committee approval.

Data Collection Tool

Turkish items of LGS were analyzed in detail. The items are made openly available within days/days following the administration of the exam. However, the data of the students who participated in the exam are not open access; information about the exam results can be accessed from the "Central Examination Report on Secondary Education Institutions" published regularly by MoNE every year. These reports include introductory information on the scope and evaluation of the exam, statistics of students who took the exam, descriptive statistics on correct answers based on subtests, central exam score distribution, and information on the relationship between exam performance and school achievement score. However, the reports do not include any item-based information.

Data Analysis

The current study analyzed 100 Turkish items from LGS 2018-2022 using the features specified in Table 1.

Table 1.
Features in the Items

Features	Variables
Test statistics	Test difficulty Test discrimination Internal consistency coefficient (KR-20)
Textual features	Sentence, word, and syllable length Readability
Content features	Whether it is visual or not Positive-negative item stem Dependence on a common stem
Purpose features	The sub-topic it aims to measure Text type (informative, narrative, poetry)

In the analysis, the contents subject to analysis in Table 1 can be listed as follows:

Examining the change in test statistics: By examining the reports published by MoNE, the results were tabulated, and the change was expressed by creating a graph.

Examining the change in textual features: In the analysis of the textual features of the items, sentence, word, and syllable length for the item stem, sentence, options, and the whole item, and the readability of the items were calculated. Notably, different readability formulas can be used in readability calculations (such as the Dale-Chall readability formula, Flesch readability formula, Coleman readability formula, and Fry readability graph). Some of these formulas have been adapted

into Turkish. For example, the Flesch readability formula was adapted into Turkish by Ateşman (1997) (Equation 1) and has gained widespread use (Baş & Yıldız, 2015; Çiftçi et al., 2007; Durukan, 2014). At this stage, calculations were performed using Python programming language, and the results were shown in tables and graphs.

$$\text{Readability value} = 198.825 - (40.175 \cdot [X1\ ] - 1) - (2.160 \cdot [X2\ ] - 2) \quad \text{(Equation 1)}$$

X1: Average length of sentences in the text in words

X2: Average length of words in the text in syllables

Although the readability of texts has been calculated, this study calculated the readability values of Turkish items for the first time. For this purpose, the process steps were introduced in detail. In the process of calculating the length and readability values of the items,

- The expressions with Roman numerals such as I and II in the items with premises are written as "one, two." If the bullet points are written as I. and II. in the item, they are edited as they are read as "first, second."
- Phrases are written as numbers, such as "5, 1980", and have been edited to read as "five, one thousand nine hundred and eighty."
- The lines in the poems are put together to form sentences.
- The statements in the options are defined as sentences.
- The words or phrases in the options are also defined as sentences. For example, the expression "Only I" is organized as "Only one," and the expression in each option is counted as a sentence.
- In the Turkish subtest of the LGS 2019 application, visual text was included in 5 items (items 4, 5, 7, 17 a, 18), and verbal expressions were shown in a table in 2 items (items 19 and 20). In the "visual reading" items, the information in the visual was described verbally. The expressions used in the description process were chosen like the expressions in other items (abstract-concrete, length, etc.). For example, visual texts were described verbally in item 4 of the 2019 exam. While the readability rate of the expressions in the item stem and the item sentence is 76.06, the readability value calculated for the whole item because of the verbal descriptions of the visuals in the options is 74.07.

Examining the change in content features: In the examination of the content features of the items, whether the items contained visuals, whether the item stem was positive or negative, and whether the items were dependent on a common stem were examined; frequency and percentage values were calculated, and the results were presented graphically. Although the readability of texts has been calculated, in this study, the readability values of Turkish items were calculated for the first time. For this purpose, the process steps were introduced in detail. In the process of calculating the length and readability values of the items;

Examining the change in the purposeful features: To examine the purpose of the items, the subtopics that the items aimed to measure, and the type of text used in the item were analyzed; frequency and percentage values were calculated, and the results were shown graphically.

Findings

1. How do LGS 2018-2022 Turkish test statistics (test difficulty, discrimination, internal consistency) change?

Based on the results of the central exam reports, the change in LGS 2018-2022 Turkish test statistics was analyzed. In the 2018 numerical report, the mean and standard deviation values for the subtests are reported based on the answers of the students who were placed, not the number of students who participated in the exam. However, the test difficulty of the subtests was calculated based on all candidates who took the test. For comparison, the average value for 2018 was calculated based on the test difficulty. The results for 2018-2022 are shown in Table 2, and the changes over the years are shown in Figure 1.

Table 2.
Changing the Test Statistics for LGS 2018-2022 Turkish Items

Booklet	Number of students	Number of items	X	S_x	Test difficulty	Test discrimination	Internal consistency (KR-20)
LGS 2018	971.657	20	12.40	----	0.62	0.56	0.84
LGS 2019	1.029.555	20	11.75	5.15	0.59	0.59	0.87
LGS 2020	1.472.088	20	10.00	4.49	0.50	0.38	0.82
LGS 2021	1.038.492	20	9.41	4.79	0.47	0.41	0.82
LGS 2022	1.031.799	20	9.22	4.61	0.46	0.39	0.81

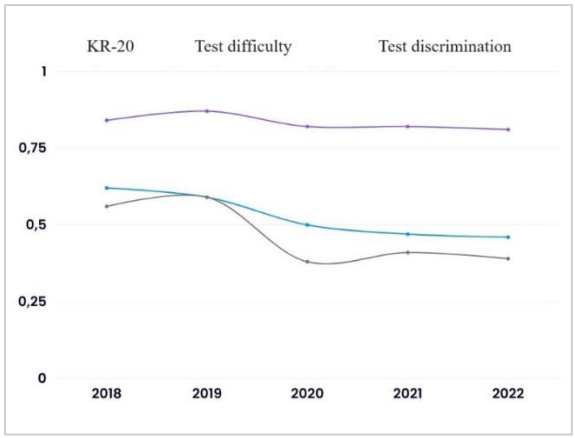

Figure 1. Changing the Test Statistics for LGS 2018-2022 Turkish Items

Average item difficulty is the average of the difficulty levels of the items in a test or subtest. This criterion provides information about the correct answer rates of the items in the test or subtest (Crocker & Algina, 1986). Figure 1 shows that the test difficulty coefficient of the LGS Turkish test has decreased over the years. In 2018, the test difficulty of the test was 0.56, while it was 0.46 in 2022. In other words, while the students who took the test in 2018 answered 56% of the Turkish items correctly on average, the students who took the test in 2022 answered 46% correctly on average. This result shows that the average correct answer rate in the LGS Turkish test has decreased.

The test discrimination coefficient is the average of the discrimination coefficients of the items in a test or subtest. This coefficient measures how much people with different achievement levels can be discriminated by the items in the test or subtest. 0.30 - 0.40 indicates adequate discrimination; 0.40 - 0.50 indicates high discrimination; and 0.50 and above indicates very high discrimination (Mertler, 2003). When Figure 1 is analyzed, it is seen that the test discriminations of the LGS Turkish test differ according to the years. While the test discrimination coefficient was close to each other and above 0.50 (very high) in LGS 2018 and 2019, this ratio decreased to 0.38 in LGS 2020. It took similar values (0.41 and 0.39 - high discrimination) in 2020-2022.

KR-20 coefficient values of 0.70 and above indicate that the internal consistency of measurement tools used in social sciences is sufficient (Cronbach, 1951; Tavakol & Dennick, 2011). When the information in Figure 1 is examined, it is seen that the internal consistency coefficients calculated for the answers given by the students to the LGS 2018-2022 Turkish items are generally similar and calculated above 0.70.

2. How do the textual features (sentence, word, syllable length, and readability) of LGS 2018-2022 Turkish test items change?

2.1. Test Length

In the process of analysing the textual features of LGS Turkish items, firstly, the number of sentences, words, and syllables in the items was calculated to determine the length of the items. The descriptive statistics calculated for the results are given in Table 3.

Table 3.
Descriptive Statistics on the Length of Items in LGS 2018-2022

Item element	Length	N of Items	Minimum	Maximum	X	Sx
All item	Sentence count	100	5.00	35.00	11.94	6.27
	Word count	100	24.00	416.00	111.15	71.53
	Syllable count	100	80.00	899.00	297.25	172.26
Stem	Sentence count	100	0.00	26.00	5.99	5.37
	Word count	100	0.00	267.00	65.63	54.11
	Syllable count	100	0.00	627.00	172.39	136.37
Question prompt	Sentence count	100	1.00	6.00	1.18	0.81
	Word count	100	5.00	54.00	11.08	8.10
	Syllable count	100	17.00	168.00	35.88	22.77
Options	Sentence count	100	4.00	24.00	4.77	2.72
	Word count	100	4.00	288.00	34.44	41.11
	Syllable count	100	4.00	600.00	88.98	100.31

When the length of a total of 100 Turkish items in the LGS in 2018-2022 is analyzed in Table 5, it is seen that there are 11.94 (±6.27) sentences, 111.15 (±71.53) words, and 297.25 (±172.26) syllables on average in the items, including item stem, item sentence, and options. It was determined that there were 5 items with direct item sentences without an item root (For example, "In which of the following sentences was a mistake made regarding the use of punctuation marks? -LGS 2018, item17). While two of these items were about spelling and punctuation, three belonged to the word meaning subtest. It was determined that the texts in the item stem had an average of 5.99 (±5.37) sentences. There were 65.63 (±54.11) words and 172.39 (±136.37) syllables on average. LGS 2018-2022 LGS Turkish items had an average of 1.18 (±0.81) sentences, 11.08 (±8.10) words, and 35.88 (±22.77) syllables. Although item sentences usually consist of 1 sentence, there are also item sentences in which other sentences are connected to the antecedents within the item sentence. The following item can exemplify this situation:

– *"According to this, in the explanations of the word "test", which of the following qualities are expected to be found in dictionaries; I. Defining the word and exemplifying its use following the definition, II. Including idioms and proverbs in which the word is used, III. It shows the meanings of some concepts with pictures. IV. Please give information about the language in which the word originated and its state in that language. V. Specify the local uses of the word. (LGS 2019; Item 18)*

There are 4.77 (±2.72) sentences, 34.44 (±41.11) words, and 88.98 (±100.31) syllables on average in all options A, B, C and D of the items. Even if there is only one word in the options, it is seen that there are at least 4 sentences since the expressions are defined as sentences. To assess the variation in item length over the years, an analysis was conducted on the word count of the item stem, item sentence, and options. The findings are illustrated in Figure 2.

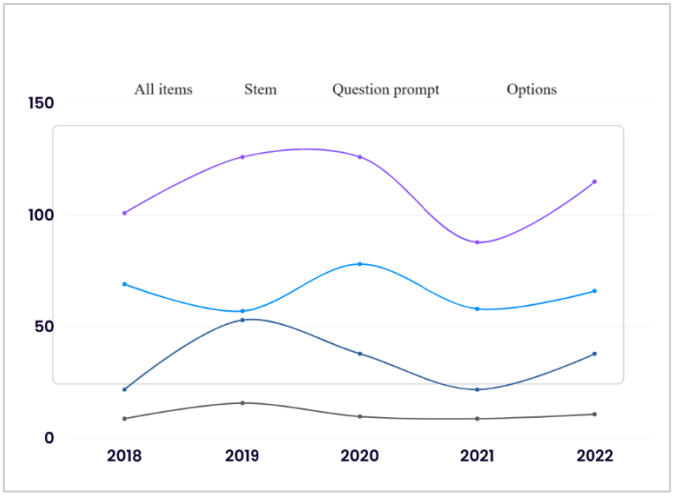

Figure 2. Changing the Length of Items in LGS 2018-2022

2.2. Readability

The calculated readability values for the Turkish items in LGS 2018-2019 are shown in Table 4 and the changes are presented in Figure 3.

Table 4.
Descriptive Statistics for the Readability Values of Items in LGS 2018-2019

Readability	Item element	N of Items	Minimum	Maximum	X	Sx
Atesman	All items	100	20.35	91.96	64.28	13.87
	Item sentence	100	0.00	77.51	39.90	15.21
	Options	100	0.00	156.04	80.93	37.26
	Item root	95	3.74	100.39	58.42	18.48

The readability values presented in Table 6 were calculated using Ateşman's (1997) readability formula. When Table 6 was analyzed, it was found that the average readability of the statements in the item stem was 58.42 (±18.48), the item sentence was 39.90 (±15.21), and the options were 80.93 (±37.26). The average readability value of all items was calculated as 64.28 (±13.87). When the calculated readability values are analyzed, it is understood that the readability of the items is generally at a medium level. It was determined that the readability of the options was easy, the item stems were medium, and the readability of the item sentences was difficult.

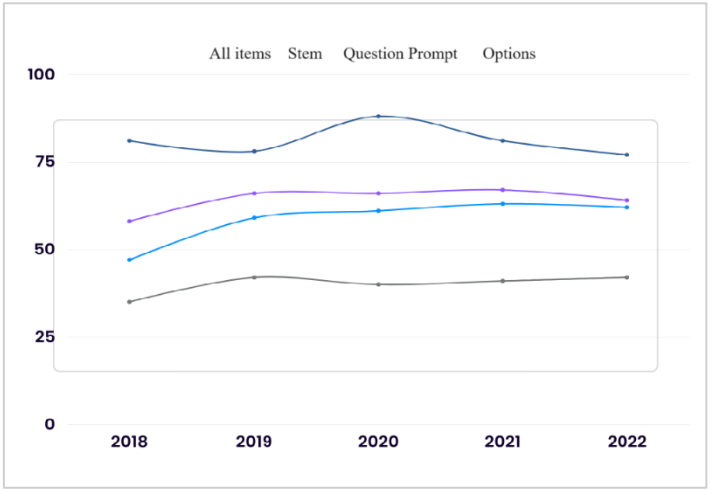

Figure 3. Changing the Readability of Items in LGS 2018-2022

Figure 3 shows that the readability of Turkish items was relatively more difficult in 2018. It is seen that the readability of the item stem, item sentence, and the whole LGS item list in 2019-2022 is consistent. Although the readability of the options was easy every year, it was determined that the readability of the options was easier in 2020 compared to other years.

2.3. Length-Readability

The Pearson correlation coefficient was calculated to determine the relationship between the length and readability of LGS 2018-2022 Turkish items. It was found that there was no significant relationship between the average word count in the items and the readability value (r=0.037; p=0.717>0.05). Relationships were also analyzed based on options. One of the principles of item writing is that the options should be equally difficult (Tekindal, 2008). For this purpose, the options should be equal in many respects, such as structure, meaning, expression, and length. In this direction, the relationship between the length and readability of the options was analyzed with Pearson correlation coefficient. The results are shown in Table 5.

Table 5.
Relationships Between the Length and Readability of the Options of the Items in LGS 2018-2022

Options	Length (number of words)				Readability			
	A	B	C	D	A	B	C	D
A	1	,976**	,969**	,982**	1	,762**	,771**	,782**
B	,976**	1	,978**	,970**	,762**	1	,774**	,873**
C	,969**	,978**	1	,973**	,771**	,774**	1	,804**
D	,982**	,970**	,973**	1	,782**	,873**	,804**	1

**p<0.001

When Table 5 is analyzed, it is seen that there is a positive and high level of correlation between the lengths of the options of the LGS 2018-2022 items (r=0.969-0.982; p<0.05). In other words, the word-based lengths of the options A, B, C, and D of the items are mainly similar. It is also understood that there are positive and high relationships between the readability of the options (r=0.771-0.873). However, the relationship between the readability of the options is lower than the word lengths. In other words, although the lengths of the options are very similar, their readability is not similar at the same rate.

3. How do the contextual features of LGS 2018-2022 Turkish test items (visuality, meaning of the item root, common stem) change?

In the study, to determine the content features of LGS 2018-2022 Turkish items, it was examined whether the items contained visual texts and poems, whether the item stem was positive or not, and whether it depended on a common stem. The calculated frequency and percentage values are given in Table 6. The change according to years is also shown in Figure 5.

Table 6.
Distribution of LGS 2018-2022 Turkish Items According to Content Features

Booklet	Content features	Yes		No	
		f	%	f	%
LGS 2018-2022	Contains visual text	19	19,0	81	81,0
	Positive item sentence	71	71,0	29	29,0
	Dependence on a common stem	7	7,0	93	93,0

When Table 8 is analyzed, it is seen that 19% of the Turkish items in LGS 2018-2022 included visual and written text. It is understood that 71% of the items consisted of positive item stems, and 29% consisted of negative item stems. 7% of the items were based on a common stem (two common stems).

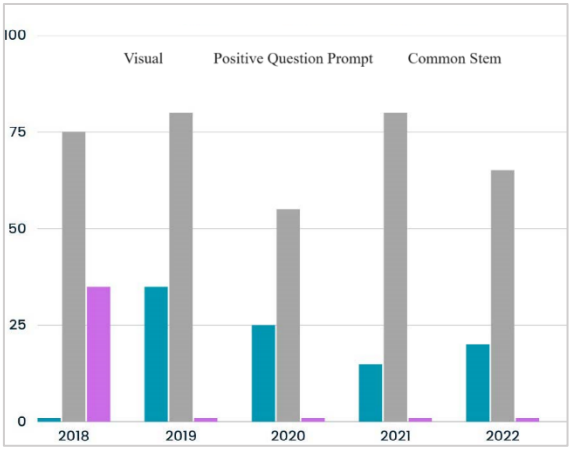

Figure 4. Changing in the Content Features of LGS 2018-2022 Turkish Items by Years

When Figure 6 is analyzed, while there were no items with visual text in 2018, 35% of the items in 2019, 25% in 2020, 15% in 2021, and 20% in 2022 included visual text. While 75% of the Turkish items in LGS 2018 were written in positive stem, 80% of the items in 2019, 55% in 2020, 80% in 2021, and 65% in 2022 were written in positive stem. In LGS, only in 2018, the first year of implementation, were items based on common stems used. A total of 7 items were asked in two common stems, 4 items based on the first common stem, and 3 items based on the second common stem. Between 2019 and 2022, no items were based on the common stem.

4. How do the purpose features (sub-topic, text type) of LGS 2018-2022 Turkish test items change?

4.1. Change in the sub-topic that the items aim to measure

The sub-subjects that LGS 2018-2022 Turkish items aimed to measure were examined in the context of curriculum outcomes. The results are presented in Table 7.

Table 7.
Distribution of LGS 2018-2022 Turkish Items According to the Sub-Subjects They Aim to Measure

Booklet	Subtopics	f	%
LGS 2018-2022	Meaning in the Word	10	10,0
	Meaning in Sentence	17	17,0
	Meaning in a Fragment	31	31,0
	Language and expression	4	4,0
	Text type	3	3,0
	Reasoning	16	16,0
	Language knowledge	11	11,0
	Spelling and punctuation	8	8,0

When Table 7 is examined, it is seen that 31% of the 100 Turkish items in LGS 2018-2022 are about meaning in passage, 17% about meaning in sentence, 16% about reasoning, 11% about grammar, 10% about meaning in word, 8% about spelling and punctuation, 4% about language and expression, and 3% about text type. In addition, 3 of the grammar items are related to expression disorder, 2 to sentence type, 1 to the elements of a sentence, 1 to roof, and 3 to structure. In addition, it is seen that the highest number of items in the LGS Turkish test in 2018-2022 belong to the topic of meaning in the passage (31%) and the lowest number of items belong to the topic of text type (3%). At the same time, the items on the meaning in the passage were the most predominant topics in all years. It is understood that the weight distribution of grammar, spelling and punctuation items is generally similar across years (this topic was not included in 2021 due to the COVID-19 pandemic). The weights of language and expression, text type, and reasoning differ between years.

4.2. Change in text type

In 13% (n=13) of the 100 Turkish items in LGS 2018-2022, it was determined that the item started with a direct item sentence or the statements in the item stem did not show a text feature. The 87 items with a textual context in the item stem were analyzed in terms of the type of text they contained, and the results are presented in Table 8.

Table 8.
Distribution of LGS 2018-2022 Turkish Items According to the Text Types

Booklet	Text type	f	%
LGS 2018-2022	Informative	63	63,0
	Narrative	3	3,0
	Poetry	1	1,0
	Informative and visual	19	19,0
	Informative and narrative	1	1,0

When Table 8 is examined, it is seen that most of the Turkish items in LGS 2018-2022 were in the informative text type. Informative text was used in 63% of the items, informative text and visual content were used in 19%, and both informative and narrative text were used together in 1% of the items. Based on all these data, 83% of the Turkish items in the exam were in the informative text type. Narrative text was used in 3% of the items, and narrative text was used together with informative text in 1%. Only 1% of the 100 Turkish items in LGS 2018-2022 utilized poetry text type. Since informative text type was used in more than 90% of the items, the change over the years is not shown graphically, but the levels of text types in LGS 2018-2022 Turkish test items according to years are as follows:

- Informative text type was used in 17 items in LGS 2018, 10 items in 2019, 13 items in 2020, 12 items in 2021, and 12 items in 2021.
- While there was no narrative text type in the LGS items in 2018 and 2019, narrative text type was used in 1 item in 2020, 1 item in 2021, and 1 item in 2021.
- While poetry was included in 1 item in LGS 2018, poetry text type was not used in the exams held between 2019-2021.
- However, in LGS 2019, informative and visual text were used together in 7 items. In LGS 2020, informative and visual text were used together in 5 items, and informative and narrative text were used together in 1 item. In LGS 2021, informative and visual text were used together in 3 items, while informative and visual text were used together in 4 items in 2022.
- In addition, it was determined that there was no text in the item stem of 2 items in LGS 2018 or the expressions were not in a context that would constitute a text.

5. How do the test statistics, textual, contextual, and purpose features of LGS 2018-2022 Turkish test items change together?

The changes in test statistics, textual features, contextual features, and purposeful features of LGS 2018-2022 Turkish test items are shown in Table 9.

Table 9.
Relationships between Psychometric and Structural Features of LGS 2018-2022 Turkish Items

Booklet	Psychometric Features ($\bar{X}$)			Textural Features ($\bar{X}$)			Content Features (%)			Purposive Features (%)		
	Test difficulty	Test discrimination	KR-20	Word count	Readability	Visual	Positive Question Prompt	Common Stem	Informative text	Comprehension	Grammar	Reasoning
LGS 2018*	0.62	0.56	0.84	100.7	58.3	0.0	75.0	35.0	85.0	70.0	25.0	5.0
LGS 2019	0.59	0.59	0.87	126.3	65.8	35.0	80.0	0.0	85.0	50.0	25.0	25.0
LGS 2020	0.50	0.38	0.82	125.9	66.3	25.0	55.0	0.0	95.0	55.0	20.0	20.0
LGS 2021	0.47	0.41	0.82	88.4	67.2	15.0	80.0	0.0	75.0	85.0	0.0	15.0
LGS 2022	0.46	0.39	0.81	114.7	63.9	20.0	65.0	0.0	75.0	65.0	25.0	10.0

The results obtained when Table 9 is analyzed stated that LGS 2018-2022 Turkish test statistics are considered together; it is seen that the test difficulty coefficient for the Turkish test was 0.62 and 0.59 in 2018 and 2019, respectively, when the test discrimination was the highest. In the years when the Turkish test was the easiest for the students on average, the discrimination values and internal consistency coefficient were also the highest. As the test difficulty decreased, test discrimination also showed a general decrease. However, it was determined that the internal consistency of the tests and the curves of change in test difficulty and discrimination were different.

The average number of words in the tests increased and decreased over the years, which is similar to the changes in test difficulty and discrimination. It was determined that the test difficulty of the items in 2019, which had the highest test length (average word count), was the highest (easiest), and the test discrimination and internal consistency were the highest. However, the change curves were not similar.

Similarly, the test with the highest proportion of items with visual text (LGS 2019) also had the highest test discrimination and difficulty. The average word count was also the lowest in the years when visual text was not included at all (LGS 2018) and when it was included the least (LGS 2021). In other words, it was determined that the length of the visual items was longer than the other items. Since the common stem item was only in 2018, it is impossible to talk about a change in this regard. In 2020, the year with the lowest number of affirmative item sentences, the average stoicism was calculated to be the lowest. While there was a slight increase in test discrimination in 2021, when the number of affirmative item sentences increased, there were slight decreases in affirmative item sentences and test discrimination in 2022. No similar change was observed for the change in test difficulty and whether the item stem was affirmative or negative. The composition of the items in the tests with affirmative-negative stems also varies by year.

In 2020, when informational text was used the most, test discrimination was the lowest. However, test discrimination was lower in 2021 and 2022, when this text type was used the least. It is seen that the changes in test difficulty and discrimination are not similar to the changes in the sub-subjects that the items aim to measure.

Conclusion, Discussion and Suggestions

This study aims to examine the structural and psychometric features of LGS Turkish items implemented since 2018 and determine their relationship. For this purpose, 100 LGS Turkish items were analyzed psychometrically in terms of test difficulty index and item difficulty index and structurally in terms of textual, content, and purpose features. As a result of the evaluation, firstly, the results obtained within the scope of the research, the discussion of these results in the literature, and finally, the recommendations are given.

Test statistics

It is understood that the test difficulty of LGS Turkish items has decreased over the years; in other words, the items have become more difficult over time. While the discrimination levels of the tests were relatively high in 2018 and 2019, they were found to be low in 2020-2022 (although they still had high discrimination).

Textural features

The number of words was chosen to compare the length of the items. It is seen that the part with the most expressions among the elements that make up the item is the item stem, then the options, and finally, the item sentence. This is expected, considering that most of the items aim to measure reading comprehension skills. Although the length of the items based on the number of words is generally similar, it is understood that the item lengths in 2018 and 2021 are shorter than the exams in other applications. This may be because LGS was implemented for the first time in 2018, and there were 7 items based on the common stem. In 2021, there was a partial shortening in the item lengths due to the suspension of education for a certain period due to the COVID-19 outbreak and the realization of online training.

The readability of the items was generally at a medium level. The readability of the options was found to be easy, the item stems were found to be medium, and the readability of the item sentences was found to be difficult. One of the reasons for this situation is that the item sentences usually consist of a single sentence and, therefore, include expressions in the question prompt. The readability of the LGS items in 2019-2022 was found to be of medium difficulty and consistent. Although the lengths of the options were very close to each other on a word basis, the readability of Option B was generally more difficult than the other options.

Content features

In 2018, the first year of implementation, there were no items with visual texts to measure visual reading skills, while the number of items with visual texts was consistent in 2019-2022. It has been determined that the distribution of positive and negative roots of Turkish items in LGS 2018 showed inconsistencies. Regarding measurement and evaluation principles, items should be written in a positive stem. Due to their nature and the small number of items in the exam, some items can be expected to be written in negative stems. For example, an item aimed at measuring the acquisition of "Determines the main idea / main emotion of the text" should be written in the positive root. However, given that a text in an item aimed at assessing the outcome 'Identifies the auxiliary ideas in the text' may contain multiple auxiliary ideas, it is impractical to create individual items for each. Consequently, the item designed to measure this outcome may adopt a negative format. Notably, in the inaugural year of LGS implementation in 2018, items based on common stems were employed. A total of 7 items were presented, with 4 items centered on the first common stem and 3 items on the second common stem. Subsequently, between 2019 and 2022, no items were framed around common stems. The absence of common stem-related items during this period, attributed to the limited item count, is considered a positive outcome.

Purposive features

In the LGS applications of 2018 and 2019, the poetry text type was incorporated into only one item. Notably, in 2019, it was discerned that the verses within the item's options were intended not for poetry interpretation but for the comprehension of visual texts. Furthermore, it was observed that no items for poetry interpretation were present in the Turkish sections of the applications from 2020 to 2022. Considering the inclusion of the outcome 'Writes poetry' in the 8th-grade curriculum, incorporating items featuring poetry becomes essential for ensuring content validity. Despite the emphasis on narrative texts in both the curriculum and textbooks, the sporadic exclusion or minimal utilization of narrative texts in LGS in certain years is inconsistent with content validity principles.

Changing together

A relatively inverse correlation was observed between test discrimination and the positivity of the item stem. As the items were written in negative stems, a decrease was observed in test discrimination. It is a necessity to analyze these descriptive results on an item basis in further studies. The test in 2019, the longest test, was determined to be the test with the easiest test difficulty and the highest discrimination. However, item-by-item examinations are necessary to determine whether the test becomes easier or harder as the item length increases. It was found that the readability of the items, the inclusion of visuals, being in positive or negative stems, and the type of text did not directly affect the items' difficulty levels. While some visual items focus on comprehension, some are aimed at measuring reasoning skills. Therefore, what the visual text measures is more important than the text itself. At the same time, it was determined that the length of the visual items was longer than the other items. However, the effect of this situation on difficulty should also be examined based on student data.

There are many studies in the literature on LGS Turkish items (Aktaş, 2022; Altun, 2021; Batur et al., 2019; Benzer, 2019; Calp & Alpkaya, 2021; Çiçek & Dilekçi 2022; Ekinci & Bal, 2019; Erden, 2020; Kanık Uysal, 2022; Kılkapan & Nacaroğlu, 2019; Ordu et al., 2021; Soysal & Güngör, 2022). For example, Calp and Alpkaya (2021) found that LGS items predicted the learning outcomes related to reading and writing skill areas in the Turkish curriculum but did not include any content related to the learning outcomes related to listening/watching and speaking skill areas. When the outcomes included were examined, it was understood that items covering reading interpretation, making

inferences, spelling, punctuation, sentence types, verbs, and expression disorders were included. It is stated that items on similar topics were asked in all LGSs from 2018 to 2022 (Aydın, 2022). Erden (2020) and Sayın and Takıl (2023) stated that Turkish teachers think that skill-based items focus on specific learning outcomes and do not cover all learning areas. Therefore, content validity in LGS exams cannot be fully ensured. These issues overlap with Benzer's (2019) study examining Turkish textbooks for LGS exams. Another related result was found in Diker Coşkun's (2013) study, which examined textbooks in terms of PISA reading criteria, and it was concluded that textbooks were insufficient in developing reading skills that require higher-order thinking skills. In Altun's (2021) study examining the theme evaluation items in the 8th-grade Turkish textbook according to PISA levels and Bloom's revised cognitive domain taxonomy, it was found that there were no items belonging to the steps of analysis, evaluation, and creation. Erden (2020) states that Turkish teachers also state that textbooks are not compatible with skill-based items and that they experience a shortage of resources. Again, the items asked of the students in Turkish exams were also examined in this context, and it was determined that there were only items for the creation step among the higher-level thinking skills according to the cognitive process dimension of Bloom's taxonomy; there were no items for the analysis and evaluation steps, and there were no items for higher level reading skills (5th and 6th level) according to PISA reading skills proficiency levels. These results indicate that Turkish course written exams are insufficient in measuring higher-level thinking skills (Kanık Uysal, 2022). In the study conducted by Büyükalan Filiz and Yıldırım (2019), the secondary school Turkish curriculum was examined according to the renewed Bloom's taxonomy. It was understood that the gains were concentrated in the comprehension and application steps in the cognitive process dimension. There were very few gains in the knowledge accumulation dimension at the metacognitive level. A similar study conducted by Çerçi (2018) determined that the learning outcomes were insufficient to represent the steps of analyzing, evaluating, and creating, which require high-level thinking skills. It was observed that the skills addressed in all test types and activities in the textbook did not show a balanced distribution in cognitive stages and were concentrated in the comprehension and analysis stages. In addition, it was concluded that the recall stage was not included in the LGS Turkish test, the application stage was not included in the LGS Turkish test and sample items, the evaluation stage was not included in the LGS Turkish test, sample items and theme evaluation items, and the creation stage was not included in the LGS Turkish test and sample items (Taşıyaran, 2022).

The content that students encounter in the Turkish course has also been compared with PISA, an international exam. It has been determined that the reading skill outcomes in the Turkish curriculum do not match the PISA content, and the outcomes that do match are concentrated at levels 1, 2, 3, and 4 (Batur & Ulutaş, 2013; Batur et al., 2019; İnce, 2016, Karabulut, 2017; Koç, 2021). When the items in Turkish textbooks were compared according to PISA reading proficiency levels, it was found that the items measured subcognitive processes (Benzer, 2019; Bozkurt et al., 2015; Yağmur, 2009), there were almost no items at the 5 and 6 levels in Turkish textbooks, the weight of the items was in the cognitive process of accessing and remembering information (Bozkurt et al., 2015) and the principles of the current reading approach were not used in reading activities (Yağmur, 2009). However, 75% of the PISA reading skill assessment framework includes metacognitive level items (Bozkurt et al., 2015). An indicator of this is the timing of graph/table reading and visual interpretation skills. This item type is also included in LGS and sample items published by the ministry. Köse and Kanık Uysal (2020) state that visual literacy, which is based on understanding and solving expressions in visual elements, is considered an essential skill in today's conditions. Success in a centralized exam is important for evaluating students' academic achievement. Turkish items in these exams are interpreted not only as an indicator of academic success but also as an output that shows that individuals who can communicate effectively with society have grown up. In this respect, learning to read is the foundation of academic success and is considered one of the most outstanding achievements in childhood (Ehri, 1995; Paris, 2005). Pressley (1998) argues that it is important for students to perceive themselves as successful readers for their academic success. It is only possible for a student to characterize himself/herself as a successful or active reader when he/she makes sense of what he/she reads, in other words, when he/she reaches meaning from the text. Research on reading emphasizes that reading is a process of making meaning (Dadandı, 2020; Temizkan, 2008,

2009; Uyar, 2015; Yılmaz, 2008). In LGS, students' realization of this meaning-making process is itemized. In this inquiry, students are expected to use their world knowledge while reaching meaning in the text they read.

The Ministry of National Education's 2023 Education Vision Document states that one of the goals is to reduce the need for competitive and qualifying exams. In this context, the central exam's purpose, content, and item types should be reconsidered, and the level of students' internalization and use of 21st-century skills should be measured. Thus, it aims to adopt a measurement and evaluation approach in which basic cognitive skills are measured (MEB, 2018a). The change in the system applied in the transition to high schools in 2018 was also influenced by this education policy. Batur and colleagues (2019) state that Turkey's central exams for transition to high school were conducted with a system that asked for memorized knowledge until 2018 and did not focus on reading comprehension and interpretation. Therefore, an inference can be made regarding the level of Turkish that students transfer to daily life in creating the content of LGS items.

In this study, the items were evaluated in terms of text type, readability, and visual usage, and an attempt was made to present a perspective by including the results of the studies in the literature on the subject. Based on the findings and results obtained, the following suggestions can be made:

- In this study, it was determined that there was no direct relationship between test length and test difficulty and that the student average was the highest in the test with the highest test length. These results should be analyzed on an item basis to prevent the perception that longer items are more difficult.
- Since exam duration remains constant, it is recommended that the average length and readability of the items should be organized in such a way that they are consistent across years.
- The readability values of LGS items should also be determined to assess the extent to which students are exposed to a reading load. It was observed that although the length of the options was the same, their readability was different. It is recommended that this issue should also be taken into consideration in item writing principles.
- To ensure content validity in LGS Turkish items, content from each text type should be included so that students can complete the process of reaching an understanding of different text types. LGS should adopt a holistic approach that includes reading and writing outcomes and listening and speaking outcomes.
- A conscious choice should be made as to which achievements can be measured with positive item stems and which with negative item stems. In this context, it is recommended to prepare a taxonomy and an item writing principle guide for constructing reading comprehension items.
- Pre-calibration studies should be conducted to ensure that the psychometric features of the items are consistent across years.
- It is recommended to publish statistics on an item-by-item basis.
- It is recommended that similar studies be conducted in exams related to different transition systems.

Limitations of the Study

In this study, changes in the structural features of the tests were analyzed based on descriptive statistics. More detailed relationship examinations can be made based on real data from students.

Acknowledgment

Ethics statement: In this study, we declare that the rules stated in the "Higher Education Institutions Scientific Research and Publication Ethics Directive" are complied with and that we do not take any of the actions based on "Actions Against Scientific Research and Publication Ethics." At the same time,

we declare that there is no conflict of interest between the authors, which all authors contribute to the study, and that all the responsibility belongs to the article authors in case of any ethical violations.

Author Contributions: Conceptualization, methodology, validation, and analysis by the first author; writing, review, and editing; supervision, and analysis by the second author.

Funding: This research received no funding.

Institutional Review Board Statement: In this research, all documents were accessed from MoNE's official website (https://raporlar.meb.gov.tr/); no participants were involved. All human research and all other activities that involve partially human research, regardless of whether there is a sponsor/supporter, must be reviewed and approved by the relevant ethics committee before data collection can begin. All research involving human participants (including interviews, questionnaires, and questionnaires) must be assessed by the relevant ethics committee, except for research that does not require ethics committee approval.

Data Availability Statement: Data generated or analyzed during this study should be available from the authors on request.

Conflict of Interest: Authors should declare no conflict of interest among authors.

References

Aktaş, E. (2022). 2018-2021 LGS Türkçe sorularının PISA okuma yeterlik düzeylerine göre değerlendirilmesi *Uluslararası Türkçe Edebiyat Kültür Eğitim Dergisi, 11*(1), 258-276.

Altun, K. (2021). *8. sınıf Türkçe ders kitabındaki tema değerlendirme sorularının PISA düzeylerine ve yenilenmiş Bloom Taksonomisi'ne göre incelenmesi* (Yayımlanmamış yüksek lisans tezi). Necmettin Erbakan Üniversitesi Eğitim Bilimleri Enstitüsü, Konya

Ateşman, E. (1997). Türkçede okunabilirliğin ölçülmesi. *Ankara Üniversitesi TÖMER Dil Dergisi*, 58, 171–174.

Aydın, İ. (2022). *Liselere geçiş sınavında (LGS) yer alan Türkçe sorularının Türkçe dersi öğretimi programı ve PISA okuma becerileri yeterlikleri bağlamında incelenmesi* (Yayımlanmamış yüksek lisans tezi). Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü, Antalya.

Baş, B. & Yıldız F. İ. (2015). Türkçe ders kitaplarındaki metinlerin okunabilirlik açısından incelenmesi.
Mersin Üniversitesi Eğitim Fakültesi Dergisi, 11 (1), 52-61.

Batur, Z., Ulutaş, M. & Beyret, T. N. (2019). 2018 LGS Türkçe sorularının PISA okuma becerileri hedefleri açısından incelenmesi. *Milli Eğitim Dergisi, 48*(1), 595-615.

Benzer, A. (2019). Türkçe ders kitaplarının PISA okuma yeterlik düzeyleri ile imtihanı. *Okuma Yazma Eğitimi Araştırmaları, 7*(2), 96-109.

Bozkurt, B., Uzun, G. L. & Lee, Y. (2015). Korece ve Türkçe ders kitaplarındaki metin sonu sorularının karşılaştırılması: PISA 2009 sonu sorularına dönük bir tartışma. *International Journal of Language Academy, 3*(4), 295-313.

Büyükalan Filiz, S. & Yıldırım, N. (2019). Ortaokul Türkçe dersi öğretim programı kazanımlarının revize edilmiş Bloom taksonomisine göre analizi. *İlköğretim Online, 18* (4), 1550-1573.

Büyüköztürk, Ş., Çakmak, E.K., Akgün, Ö.E, Karadeniz, Ş. ve Demirel, F. (2021). *Bilimsel araştırma yöntemleri*. PegemA Akademi, Ankara.

Calp, M. & Alpkaya, C. (2021). LGS Türkçe sorularının Türkçe dersi öğretim programı kazanımlarına uygunluğu üzerine bir çalışma. *Ana Dili Eğitimi Dergisi*, *9*(2), 632-654.

Çerçi, A. (2018). Türkçe dersi öğretim programı kazanımlarının (5, 6, 7, 8. sınıf) yenilenen Bloom taksonomisine göre incelenmesi. *Okuma Yazma Eğitimi Araştırmaları, 6*(2), 70-81.

Çiçek, S. & Dilekçi, A. (2022). LGS Türkçe sorularının PISA okuma ölçütlerine göre değerlendirilmesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 42(2), 1535-1553.

Çiftçi, Ö, Çeçen, M. A. & Melanlıoğlu, D. (2007). Altıncı sınıf Türkçe ders kitaplarındaki metinlerin okunabilirlik açısından değerlendirilmesi. *Elektronik Sosyal Bilimler Dergisi*, 6 (22), 206-219.

Dadandı, U. P. (2020). *Okuduğunu anlamanın bilişsel-motivasyonel faktörlerle açıklanması ve okuduğunu anlama güçlüklerine yönelik bir müdahale program* (Yayımlanmamış Doktora Tezi). Trabzon Üniversitesi Lisansüstü Eğitim Enstitüsü, Trabzon.

Davis, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika*, 9 (3), 185-197.

Diker Coşkun, Y. (2013). Türkçe ders kitaplarının PISA sınavı okuma ölçütleri açısından incelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, (26), 22-43.

Durukan, E. (2014). Metinlerin okunabilirlik düzeyleri ile öğrencilerin okuma becerileri arasındaki ilişki. *Ana Dili Eğitimi Dergisi*, 2 (2), 68-76.

Ehri, L.C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*, 116–125.

Ekinci, O. & Bal, A. P. (2019) 2018 yılı liseye geçiş sınavı (LGS) matematik sorularının öğrenme alanları ve yenilenmiş Bloom taksonomisi bağlamında değerlendirilmesi. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, *7* (3).

Erden, B. (2020). Türkçe, matematik ve fen bilimleri dersi beceri temelli sorularına ilişkin öğretmen görüşleri. AJER, 5 (2), 270-292.

Frankel, J. R. & Wallen, N. E. (2009). *How to design and evaluate research in education*. (7th Ed.). New York, NY: McGraw-Hill International Edition.

Gabb, Sally. (2000). From talk to print: Preparing students to read with ease. *Field Notes*, *10*(2).

Grabe, William. (1991). Current developments in second language reading research. *TESOL Quarterly*. 25 (3), 375-406.

Guthrie, J. T. & Scafiddi, N. T. (2004). Reading comprehension for information text: Theoretical meanings, developmental patterns, and benchmarks for instruction. In J. T. Guthrie, A. Wigfield & K. C. Perencevich (Eds.), *Motivating reading comprehension: Concept- Oriented Reading Instruction* (ss. 225–248). Mahwah, NJ: Erlbaum.

Hiebert, E. H. (1983). An examination of ability grouping for reading instruction. *Reading Research Quarterly*, 18, 231-255.

Januarty, R. & Nima, H. N. A. (2018). Energizing students' reading comprehension through multimodal texts. *International Journal of Language Education*, 2(2), 14–22.

Jiang, X., Sawaki, Y. & Sabatini, J. (2012). Word reading efficiency and oral reading fluency in ESL reading comprehension. *Reading Psychology*, *33*, 323–349.

Kanık Uysal, P. (2022). Türkçe dersi sınavlarında yer alan soruların üst düzey düşünme becerileri açısından incelenmesi. *Ana Dili Eğitimi Dergisi, 10* (1), 136-156.

Kızkapan, O. & Nacaroğlu, O. (2019). Fen bilimleri öğretmenlerinin merkezi sınavlara (LGS) ilişkin görüşleri. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi, 9*(2), 701–719.

Köse, S. & Kanık Uysal, P. (2020, Ekim). Türkiye'de görsel okuma ve görsel sunu alanında yapılan çalışmaların betimsel içerik analizi. XII. Uluslararası Dünya Dili Türkçe Sempozyumu'nda sunulan bildiri, Ankara.

Linacre, J. M. (1993). Rasch-based generalizability theory. *Rasch Measurement Transaction*, 7(1), 283-284.

MEB (2018a) *2023 eğitim vizyonu*. http://2023vizyonu.meb.gov.tr/doc/2023_EGITIM_VIZYONU.pdf, web adresinden 15.11.2022 tarihinde indirilmiştir.

MEB (2018b). 2018 Liselere geçiş sistemi (LGS): merkezî sınavla yerleşen öğrencilerin performansı (Eğitim analiz ve değerlendirme raporları serisi 3). Ankara: MEB Yayınları.https://www.meb.gov.tr/meb_iys_dosyalar/2018_12/17094056_2018_lgs_rapor.pdf

MEB (2019a). *Türkçe dersi öğretim program (1-8. Sınıflar için)*. Ankara: MEB

MEB (2019b). 2019 Ortaöğretim kurumlarına ilişkin merkezi sınav (Eğitim analiz ve değerlendirme raporları serisi 7). Ankara: MEB Yayınları.

MEB (2021). *2021 Liselere geçiş sistemi merkezî sınavla yerleşen öğrencilerin performansı*. Ankara: MEB Yayınları.

MEB (2022). *2022 Ortaöğretim kurumlarına ilişkin merkezî sınav.* Ankara: MEB Yayınları.

Melanlıoğlu, D. (2021). Okumada anlam kurma süreci ve PISA sınavı. (Editörler: Seçkin Aydın ve Ömer Yahşi). *PISA ve okuma kültürü*. Ankara: Nobel Yayıncılık. s.159-197.

Myford, C. M., & Wolfe, E. W. (2003). Detecting and measuring rater effects using many-facet Rasch measurement: Part I. *Journal of Applied Measurement*, 4(4), 386-422.

OECD (2019). *PISA 2018 assessment and analytical framework*. Paris: OECD Publishing.

Orasanu, Judith (Ed.). (1986). *Reading comprehension: From research to practice*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ordu, H. H., Engin, A. S., & Topçuoğlu Ü. F. (2021). Comparison of Turkish lesson sample items for LGS published by MEB in 2018-2019 academic year and 2019 LGS Turkish items. *SDU International Journal of Educational Studies, 8*(2), 102-113.

Özdemir, E. (2011). *Eleştirel okuma*. Ankara: Bilgi Yayınevi.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, 40 (2), 184-202.

Perfetti, C. A. & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67*(4), 461–469.

Pressley, M. (1998). Comprehension strategies instruction. I. Osborn & F. Lehr (Ed.), *Literacy for all: Issues in teaching and learning*. New York: Guilf, 113-133.

Sayın, A., & Takıl, N. B. (2023). Üst Düzey Düşünme Becerilerinin Ölçülmesine Yönelik Soruların Hazırlanmasında Karşılaşılan Zorluklar ve Çözüm Önerileri: Okuduğunu Anlama. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi, 43*(1), 531-561.

Sevgi, S. & Karakaya, M. (2021). Ortaokul öğrencilerinin okuma alışkanlığı ve problem çözme becerisinin incelenmesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 41 (2), 1203-1225.

Singer, H. (1978). Active comprehension: From answering to asking items. *The Reading Teacher, 31,* 901–908.

Soysal, T. & Güngör, H. (2022). LGS'de dil bilgisi sorularının gerekliliği üzerine öğretmen görüşleri. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi* 41(1), 275-310

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16,* 32-71.

Şahin, M. (2022). *Akıcı okuma ve temel işlem akıcılığına yönelik etkinliklerin ilkokul öğrencilerinin matematik problemi çözme başarısına etkisi (Yayımlanmamış Doktora Tezi).* İstanbul Üniversitesi-Cerrahpaşa Lisansüstü Eğitim Enstitüsü, İstanbul.

Taşıyaran, M. (2022). *LGS Türkçe sorularının öğretim programı, ders kitabı ve bilişsel alan sınıflaması açılarından incelenmesi (Yayımlanmamış yüksek lisans tezi).* Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Tekindal, S. (2008). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.

Temizkan, M. (2008). Bilişsel okuma stratejilerinin Türkçe derslerinde bilgiye dayalı metinlerin okuduğunu anlama üzerindeki etkisi. *Gazi Eğitim Fakültesi Dergisi*, 28(2), 129-148.

Temizkan, M. (2009). İlköğretim öğrencilerinin okuduğunu anlama düzeyleri ile yazılı anlatım beceri düzeylerinin karşılaştırılması. *Milli Eğitim*, 183, 42-66.

Uluyol, Ç. & Eryılmaz, S. (2015). 21. yüzyıl becerileri ışığında FATİH projesi değerlendirmesi. *Gazi Eğitim Fakültesi Dergisi*, 35(2), 209-229.

Uyar, Y. (2015). *Öz düzenlemeye dayalı okuma becerisinin geliştirilmesi ve anlamaya etkisi* (Yayınlanmamış doktora tezi). Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Vural, C. (2020). *Son 10 yılda yapılan liselere giriş sınavlarında (SBS, TEOG ve LGS) yer alan Türkçe dersi sorularının yenilenmiş Bloom taksonomisi açısından değerlendirilmesi (Yayımlanmamış yüksel lisans tezi).* Gaziantep Üniversitesi Eğitim Bilimleri Enstitüsü, Gaziantep.

Wigfield, A. ve Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, 89, 420 – 432.

Yılmaz, M. (2008). Türkçede okuduğunu anlama becerilerini geliştirme yolları. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 5(9), 131-139.