

Examination of Teachers' Opinions on The Concept of New Generation Item According to Some Variables

Eray Selçuk¹ Ferhat Karakaya² Mehmet Yılmaz³

To cite this article:

Selçuk, Karakaya & Yılmaz (2024). Examination of Teachers' Opinions on The Concept of New Generation Item According to Some Variables. *e-Kafkas Journal of Educational Research*, 11, 33-47. doi:10.30900/kafkasegt.1435336

Research article

Received: 11.02.2024

Accepted: 17.04.2024

Abstract

With the replacement of TEOG in 2018 by the LGS, as known a new concept has emerged in the measurement and evaluation. This concept is called the new generation item. This research aims to determine teachers' opinions about new-generation items. This study was designed with a mixed-method strategy. Chi-square analysis was applied to the quantitatively collected data and content analysis was applied to the qualitatively collected data. A total of 192 teachers from 23 different fields participated in the study. The research data were collected using a semi-structured questionnaire form prepared by the researchers. 75 (39 %) of the teachers stated that there is no concept of a new generation item, and 117 (71 %) of them stated that there is such a concept. It was found that there was no statistically significant correlation between the branch of the teachers, their professional time and whether they attended any course/seminar/workshop and their answers about the existence of the concept of new generation items. Seven themes were revealed in the qualitative part, where teachers' views on the concept of new generation items were taken. These themes are the Formal view, Skill-based, Taxonomic view, Reading comprehension-based, Reconstructing knowledge, Daily life skills-based, and New view. According to these themes, new generation items predominantly consist of formal reading comprehension and higher-order thinking skills.

Keywords: New generation item, skill-based item, test development, item writing, measurement and evaluation

¹  Corresponding Author, erayselcuk84@gmail.com, Ministry of National Education

²  Yozgat Bozok University, Faculty of Education

³  Gazi University, Gazi Faculty of Education

Introduction

Measurement and evaluation procedures are utilized at all stages of education and training to obtain feedback for accountability in education. These procedures are also a control system to determine the quality of education and training. Accordingly, valid and reliable measurement and evaluation tools should be used for the process to function correctly (Miller, Linn & Gronlund, 2009). The first step in the measurement and evaluation procedure is to define the trait to be measured and to carry out the measurement operation. The priority for this is to have a measurement rule or a measurement tool. With the use of valid and reliable measurement tools, various data about individuals are obtained. The results obtained from measurement tools are used to make judgments about the measured traits of individuals (Gelbal, 1994).

Measurement in the education and training system is carried out through tests to determine the outcomes of the curriculum or the students' achievement standards. Tests used in education are measurement tools used to determine students' achievement levels or measure the teaching process's efficiency (Brookhart, 2010; McMillan, 2014). In a way, tests are tools for determining the traits of an individual under standard conditions (Anastasi & Urbina, 1997). These tests are designed to measure students' level of proficiency in a specific content or ability. Test development is the process of developing and creating tests that are used to measure and assess a knowledge area or topic. According to Thorndike (1982), several activities (such as defining the purpose and scope, writing test items, piloting, and item analysis) are applied successively in the test development process. One of the most essential steps in these processes is writing the items that make up the test. The item, which is part of a test or assessment tool, i.e., the most basic unit of observation (Osterlind, 1990; Haladyna, 2004), usually refers to a specific item or statement used to assess an individual's traits, abilities, or attitudes (Anastasi & Urbina, 1997). The item writing procedure involves specific guidelines depending on the item type and requires proficiency. While preparing the item, the item writer should have accurate scientific knowledge about the behaviour to be measured, should know the group to be measured well, should define the behaviour to be measured clearly, and should have sufficient measurement and evaluation knowledge about the type of item used (Selçuk, 2022). There are many exams in various fields in many institutions in our country, especially SSPC [OSYM] and MoNE [MEB], and these exams contain many sub-tests. In a country with so many exams, test development and item writing are increasingly important areas of specialization. The exam-oriented approach that dominates the education system (Taşdemir, 2015; Ural, 2016) encourages many private publishing companies to produce test books in this field. For this reason, this gap in the field is tried to be filled by many educators who have or have not received item writing preparation training.

The High School Entrance Examination (LGS), seen as a high-stakes exam in our country, is often brought to the spotlight and discussed, mainly due to its content. In the 2023 Education Vision prepared and published by the Ministry of National Education (MoNE) in October 2018, the "Competency-Based Assessment and Evaluation System" is mentioned. This system states that the assessment of students' cognitive skills, such as reasoning, creative thinking, critical thinking, reflective thinking, lateral thinking, interpretation, prediction, and so on, will be emphasized. An approach that eliminates the need for processes such as memorizing formulas and storing knowledge will be displayed (MoNE, 2018). In this context, MoNE published "Skill-Based Example Items" every month before the exam. Recently, MoNE published new measurement and evaluation legislation. In the measurement and evaluation legislation published in the Turkish Government's Legal Gazette dated 09/09/2023 and no. 32304, "formative measurement and evaluation" was emphasized, and decisions were made to use different types of items, mainly open-ended (MoNE, 2023). At the same time, the legislation emphasizes that the concept of skill-based assessment and evaluation will continue. Many teachers and publishing companies refer to skill-based questions as "new generation questions or items." This concept has recently appeared in the LGS and is defined differently by specialists.

These items, which are very similar to the items asked in international exams, are often referred to as new generation items, skill-based items, or context-based items in the education system (Erden, 2020; Kertil et al., 2021). Çaldıran and Özkan (2022) defined new generation items as complex items that require students to comprehend what they read, make inferences, solve problems, analyze, think

critically, and use scientific processing skills. Karakeçe (2021) referred to such new generation items that develop higher-order thinking skills in students and enable them to think critically as "skill-based items". Sanca et al., (2021) stated that new generation items aim to provide students with high-level skills. Yiğit et al., (2022) noted that new generation items aim to measure students' higher-order thinking skills such as reasoning, inferring, problem solving, analyzing, critical and creative thinking. Şahan and Şahin (2023) defined new generation items as questions that help students understand, construct, interpret, and analyze information associated with real life and aim to move students away from memorization and help them think. They also stated that the new generation items are problems created with real-life shapes and scenarios, which the MoNE has defined as skill-based since 2018. In general, it is seen that the new generation items measure high-level thinking skills, are related to daily life, and are presented with a wide range of content.

Reviewing the Turkish literature, context-based and skill-based items were prepared and used (Elmas & Eryılmaz, 2015; Kabuklu et al., 2019; Şan & İlhan, 2022; Türkel, 2022), context-based items were compared with traditional items or PISA items (Sak, 2018; Tekbıyık & Akdeniz, 2010; Ünal, 2019), students' solving processes of context-based and new generation items (Nasırlıel, 2020; İlkörücü & Altaş, 2022), the opinions of teachers and students on skill-based and LGS items (Erden, 2020; Kablan & Bozkuş, 2021; Kertil et al. 2021; Çaldıran & Özkan, 2022; Karabulut et al., 2022; Ceylan & Orhan, 2023; Kaya & Şahin, 2023), and the development of attitude and perception scales for new generation items (Kılcan, 2021; Yiğit et al., 2022). The item types examined in these studies are mostly called new generation items. It was observed that only qualitative data were collected in the studies. Examining qualitative data and various variables will provide deeper insights into these studies. When the literature on assessment and evaluation in other countries was reviewed, no such concept (new generation item) was found, especially in Web of Science (WoS) and Scopus databases. This is because there is no such type among the commonly accepted types of substances globally. This study investigated the reasons for this concept's appearance and common use in our country. Accordingly, this study examines teachers' views on new generation items according to various variables. For this purpose, answers to the following research questions were sought:

1. Is there a significant difference in whether there is a "new generation item" concept according to the fields of teachers?
2. Is there a significant difference in whether there is a "new generation item" regarding teachers' occupational years?
3. Is there a significant difference in whether teachers have a "new generation item" concept according to whether they attended any course, seminar, or workshop related to measurement and evaluation?
4. What are teachers' opinions on the concept of "new generation items"?

This research is distinguished from other research because teachers' views are examined in terms of various variables; qualitative and quantitative data are collected simultaneously. In this study, the reason for determining teachers' fields, occupational years, and whether or not they have taken any measurement and evaluation courses as quantitative variables is that these variables directly or indirectly affect the concept of "new generation item." Because the concept of "new generation item" is used mainly by mathematics and science teachers in the field. This is because these two disciplines are more important than others in LGS. Therefore, it is thought that the teachers' field will have a different concept. Likewise, it is believed that this concept will differ in terms of the professional duration of teaching. Because this concept is used more frequently by teachers new to the occupational period. Accordingly, does the increased occupational duration change the thoughts about this concept? Thirdly, the reason for asking whether any measurement and evaluation courses were taken is that it is known that this concept has no equivalent in the measurement and evaluation literature. Experts frequently state that there is no such concept in such courses. Therefore, it is thought that there will be a difference between those who attended such a course and those who did not.

A review of the literature on measurement and evaluation reveals that there is no such concept as new generation item. However, this concept has recently been frequently used in test development and item writing processes, and many studies have been conducted on it. In this study, if there is such a concept,

how teachers define it was emphasized. Based on these definitions, it was tried to explain with scientific evidence that nothing is new. This study will contribute to the literature by showing that teachers misuse a non-scientific concept.

Method

This section includes information about the research model, study group, process, data collection, data analysis, and validity and reliability in the qualitative part of the research.

Research Design

This research is a mixed-method research design. Mixed-method research is a research method used in the fields of education, health, social and behavioural sciences where the researcher collects both quantitative and qualitative data to understand the research purpose, combines the two data sets, and then draws conclusions using the superior advantages of these two data sets (Creswell, 2021). The research uses a concurrent nested mixed method design. Concurrent nested mixed method design is a research design quantitative and qualitative data are collected and analyzed at the same time. However, the emphasis is usually on either quantitative or qualitative data. In these designs, one type of data is embedded in the other, so less attention is paid to the embedded data type. One reason is that the less important type of data is being used to help answer an entirely different question or set of questions. Combining data is usually done at the data analysis stage. This design is useful when it is desirable to gain a broad view of the studied topic and to survey different groups or levels within a study. (Creswell, 2021). The mixed method was preferred in this research because it is thought that by using qualitative and quantitative data together, the weakness of each data set will be strengthened by the other. The mutual results will contribute to the validity of the research by complementing each other.

Sample

The study group consisted of 192 teachers working in various cities of Turkey, from 23 different disciplines and with different occupational durations, who participated or did not participate in any measurement and evaluation seminar/course/workshop. Descriptive statistics of the study group is given in Table 1, Table 2 and Table 3. Descriptive statistics of teachers according to their fields of study is given in Table 1.

Table 1.

The Frequency and Percentage Frequency Values According to Teachers' Branches

Branches	N	%	Branches	N	%
Physical Education	1	0.5	Mathematics	21	10.9
ICT Technologies	2	1.0	Music	1	0.5
Biology	8	4.2	Special Education	2	1.0
Geography	4	2.1	Psychological counseling and Guidance	3	1.6
Religious Culture and Ethics	4	2.1	Primary School Education	24	12.5
Philosophy	3	1.6	Social Sciences	9	4.7
Science Education	43	22.4	History	1	0.5
Physics	4	2.1	Technology and Design	6	3.1
Arts	1	0.5	Turkish Language and Literature	7	3.6
IHL Vocational Lessons	4	2.1	Turkish	18	9.4
Elementary Mathematics	9	4.7	Foreign Languages	12	6.3
Chemistry	5	2.6	Sum	192	100

In Table 1, science teachers participated in the research the most with 43 (22.4 %) participants, while physical education, visual arts, music and history teachers participated the least with one (0.5 %) each. Descriptive statistics of the teachers according to their occupational duration is given in Table 2.

Table 2.

The Frequency and Percentage Frequency and Percentage Frequency Values According to Teachers' Occupational Periods

Occupational Periods	N	%
0-5 years	16	8.3
6-10 years	49	25.5
11-15 years	52	27.1

Table 2 continuing

16-20 years	42	21.9
21 or above years	33	17.2
Sum	192	100

In Table 2, 16 (8.3 %) between 0-5 years, 49 (25.5 %) between 6-10 years, 52 (27.1 %) between 11-15 years, 42 (21.9 %) between 16-20 years, and 33 (17.2 %) between 21 years and above. Descriptive statistics about whether the teachers participated in any courses, seminars or workshops related to measurement and evaluation is given in Table 3.

Table 3.

The Frequency and Percentage Frequency Values According to Teachers' Participation in Courses/Seminars/Workshops

Participation in Courses/Seminars/Workshops	N	%
Yes	142	74
No	50	26
Sum	192	100

In Table 3, it is seen that 142 (74 %) of the teachers participated in any course, seminar or workshop related to measurement and evaluation, while 50 (26 %) of the teachers did not. In the following process, the data collection phase was started.

Process

Ethical approval was obtained from Yozgat Bozok University Social and Human Sciences Ethics Committee to conduct the research with the decision dated 20/04/2022 and no. 32/40. At the same time, the necessary authorization was obtained from the Ministry of National Education to conduct the online application to teachers all over Turkey. Data were collected through Google Forms via the link <https://forms.gle/y2p3orKYWVVEPDVo7>. Teachers participated in the data collection phase voluntarily. The participants were assured that the data were collected only for academic purposes.

Data Collection

The data were collected through a semi-structured questionnaire form developed by the researchers. This form consists of five questions. The first four questions constitute the quantitative data of the research, and the fifth question constitutes the qualitative data. The fifth question is opened according to the answer given to the fourth question. This question is, "Is there such a thing (concept, item type, etc.) as "New Generation Item?". If the participant answered "No" to this question, the data collection phase ended; if the participant answered "Yes," the fifth question was opened. The fifth question of the form is "What does the New Generation Item mean to you?". After the data were collected, the quantitative and qualitative data were analyzed.

Data Analysis

The data were analyzed in two steps. In the first step, the Chi-square independence test was performed on the variables mentioned for the first three questions of the research. The chi-square test of independence is a statistical method used to determine whether there is a significant relationship between two categorical variables. This test involves comparing observed frequencies in a contingency table with expected frequencies under the assumption of independence between two variables (Koch, 1982). The Chi-square statistic is then calculated, and the difference between the observed and expected frequencies is measured. A significant difference between the observed and expected frequencies indicates a relationship between the two variables (Agresti, 2018). For the Chi-square independence test to give a robust result, the lowest expected frequency in each cell should be 5 or more. Some authors state that at least 80% of the cells should have an expected frequency of 5 or more. In cases where this is not the desired frequency, Fisher's Exact Likelihood Test results should be reported in addition to the Chi-square values. In addition, when there are two categories in each variable, i.e., in a 2 x 2 Chi-square design, Yates Correction for Continuity is required (Pallant, 2015). Therefore, Fisher's Exact Likelihood Test was performed for the first research problem, Chi-square independence test for the second problem, and Yates Correction for Continuity as well as Chi-square independence test for the third research problem.

In the second step, content analysis was performed on the qualitative data for the fourth problem of the research. Content analysis is the process of systematically coding the collected data and collecting them under specified themes (Büyüköztürk et al., 2013). There are four phases in content analysis. These are (1) Coding the data, (2) Creating themes of the coded data, (3) Organizing the codes and themes, and (4) Creating, interpreting, and reporting the findings (Yıldırım & Şimşek, 2021).

Validity and Reliability in Research

In the qualitative section of the research, the similarity ratio, which is defined as internal consistency in the Miles and Huberman formula for reliability and defined as the agreement between the coders, was calculated with the following equation: $\Delta = C \div (C + \partial) \times 100$: $\Delta = C \div (C + \partial) \times 100$. In the equation, Δ : Reliability coefficient, C: The number of topics/terms on which there is agreement, ∂ : Represents the number of topics/terms on which there is disagreement. According to the coding check that provides internal consistency, it is expected that the agreement between the coders should be at least 80% (Miles & Huberman, 1994). According to this method, the codes and themes created by the responsible author were given to the other two researchers, who were asked to perform recoding. Accordingly, the reliability between the three researchers was calculated as .87.

In the qualitative part of the study, the validity study was conducted as internal (credibility). Triangulation/triangulation method was used for internal validity. In this method, multiple methods, multiple data sources and a large number of participants are utilized in data collection (Creswell, 2020; Merriam, 2013). In this research, internal validity was ensured by working with 192 participants from 23 different disciplines. At the same time, the data obtained from the quantitative part of the research were supported by the data obtained from the qualitative part. Hence, it was meticulously examined whether the participants were accurate in their responses.

Findings

The findings of the analysis conducted in response to the research questions are given under sub-headings.

Is there a significant difference, according to the branches of the teachers, about whether there is a new generation item concept or not?

The results of the Chi-square independence test on whether there is a new generation item concept according to the branches of the teachers are given in Table 4.

Table 4.

Results of Chi-Square Independence Test Regarding Whether Teachers Have a New Generation Item Concept According to Their Branches

Branches	New generation item		Sum	χ^2	df	p^*	Fisher' p^*
	Yes	No					
Physical Education	1	0	1				
ICT Technologies	1	1	2				
Biology	4	4	8				
Geography	3	1	4				
Religious Culture and Ethics	4	0	4				
Philosophy	2	1	3				
Science Education	20	23	43				
Physics	2	2	4				
Arts	1	0	1				
IHL Vocational Lessons	3	1	4	22.830	22	.411	.372
Elementary Mathematics	4	5	9				
Chemistry	2	3	5				
Mathematics	12	9	21				
Music	0	1	1				
Special Education	2	0	2				
Psychological counseling and Guidance	2	1	3				
Primary School Education	17	7	24				
Social Sciences	5	4	9				
History	0	1	1				

Table 4 continuing

Technology and Design	6	0	6
Turkish Language and Literature	4	3	7
Turkish	14	4	18
Foreign Languages	8	4	12
Sum	117	75	192

* $p < .05$

In Table 4, it was found that the difference observed in the opinions of teachers from different branches on whether the concept of new generation item exists was not significant ($\chi^2_{(22, 192)}=22.830$; $p > .05$). This means there is no relationship between teachers' branches and their views on the concept of new generation item.

Is there a significant difference in whether there is a new generation item concept regarding teachers' occupational years?

The results of the Chi-square independence test on whether there is a new generation item concept according to the occupational years of the teachers are given in Table 5.

Table 5.

The Results of the Chi-Square Independence Test for the Presence of the New Generation Item Concept According to the Occupational Periods of the Teachers

Occupational periods	New generation item		Sum	χ^2	df	p^*
	Yes	No				
0-5 years	13	3	16	4.749	4	.314
6-10 years	28	21	49			
11-15 years	28	24	52			
16-20 years	28	14	42			
21 or above years	20	13	33			
Sum	117	75	192			

* $p < .05$

In Table 5, it was found that the difference observed in the opinions of teachers with different occupational periods on whether there is a new generation item concept is not significant ($\chi^2_{(4, 192)}=4.749$; $p > .05$). This means that there is no relationship between teachers' occupational periods and their views on the concept of new generation item.

Is there a significant difference between the teachers' participation in any course, seminar, or workshop related to measurement and evaluation and whether they have a new generation item concept or not?

The results of the chi-square independence test on whether there is a new generation item concept according to whether the teachers attended a course, seminar, or workshop on measurement and evaluation are given in Table 6.

Table 6.

The Results of Chi-Square Independence Test Regarding Whether Teachers Have a New Generation Item Concept According to Whether They Participated in a Course, Seminar or Workshop Related to Measurement and Evaluation

Participation in Courses/Seminars/Workshops	New generation item		Sum	χ^2	df	p^*
	Yes	No				
Yes	82	60	142	1.846	1	.174
No	35	15	50			
Sum	117	75	192			

* $p < .05$

In Table 6, it was found that the difference observed in the opinions of the teachers who attended a course, seminar, or workshop in the field of measurement and evaluation regarding the existence of the concept of new generation item was not significant ($\chi^2_{(1, 192)}=1.846$; $p > .05$). This means that there is no relationship between teachers' participation in any course, seminar or workshop in the field of measurement and evaluation and their views on the concept of new generation item.

Which themes are the teachers' opinions on the concept of new generation items?

The results of the content analysis of the opinions of 117 teachers who stated that there is a concept called new generation item are given in Table 7.

Table 7.

The Results of Content Analysis of Teachers' Opinions on the New Generation Item Concept

Themes	Subthemes	Codes	N	Teachers
Formal view	Text	Long text, long sentences	3	T2, T44, T130
	Visual	Graphic, Infographic, Image, Table, Figure	4	T27, T66, T109, T131
	Item type	Options, More than one topic, Meandering, Narrativized	2	T6, T178
Skill-based		Stem, Steam, Learning outcome based, Multidisciplinary	6	T3, T24, T45, T62, T104, T132
Taxonomic view	Bloom taxonomy	Comprehension, application, analysis, synthesis, evaluation, Bloom, upper stage, upper level	13	T4, T47, T60, T64, T72, T79, T89, T117, T119, T152, T157, T172, T177,
	High level skills	Skills, thinking, mind-straining, inferencing, algorithm understanding	12	T17, T18, T28, T55, T78, T85, T91, T96, T136, T162, T165, T192
	21st century skills	Millennium, 21st century	3	T43, T48, T188
	Learning outcome based	Multiple learning outcomes, outcome-item interaction, using more than one piece of knowledge, distinguishing between those who know and those who don't know	4	T54, T100, T140, T146
Reading comprehension-based		Understanding, comprehension, rapid reading, visual reading, transferring, interpreting, reading skills	19	T15, T39, T44, T102, T103, T105, T110, T118, T121, T125, T126, T135, T142, T145, T158, T159, T163, T168, T180
Reconstructing knowledge	Use of knowledge	Establishing relations, different from memorization, creating connections, testing of knowledge, prior knowledge, what is given to solve the problem	13	T9, T10, T19, T28, T34, T51, T67, T71, T73, T74, T87, T95, T164
	Interpretation of knowledge	Interpreting, explaining, inquiring, not asking for direct knowledge, evaluating knowledge, analyzing knowledge	15	T9, T11, T13, T26, T86, T98, T99, T127, T133, T138, T141, T149, T154, T160, T170
	Thinking differently	Reasoning, deduction, comprehension, progressive thinking	5	T69, T127, T137, T144, T171
Daily life skills-based	The actual world and daily life	Experiencing by doing, real life, outdoor life, daily life skills	10	T21, T31, T42, T50, T82, T112, T166, T169, T175, T179
	Gain-oriented	Multidisciplinary, Multilearning outcomes, Multibehaviour	6	T52, T68, T76, T107, T124, T186
New view		Adapting to the times, artificial intelligence assisted, next generation human	2	T22, T155
Sum				117

In Table 7, it is seen that the new generation item concept consists of seven different themes as a result of the opinions received from 117 teachers. These themes are Formal view (9), Skill-based (6), Taxonomic view (32), Reading comprehension-based (19), Reconstructing knowledge (33), Daily life skills-based (16), and New view (2).

The theme of Formal view consists of Text (2), Visual (4), and Item type (3) subthemes. Teachers stated that this concept consists of long texts and sentences in the Text subtheme. Some expressions related to teachers' views on this subtheme are given below: "Long text" (T2), "It contains slightly longer sentences and visuals." (T44).

In the Visual subtheme, teachers stated that this concept consists of visuals such as tables, figures, and graphics. Some expressions related to teachers' views on this subtheme are given below: "A Graphic, an infographic and a table is an item type that also measures visual reading skills." (T27), "The item includes information about the content as well as visuals, tables or graphics." (T109)

In the Item type subtheme, teachers stated that the options or the item's text differed in format. They also emphasized that the text part of the item was meandering and consisted of narratives. Some expressions related to teachers' views on this subtheme are given below:

...items with edited options... (T6)

The new generation items are a bit too meandering; as far as I observe, it is a style that keeps students away from the lessons; when they see the item, they do not even want to solve it. I think it should be shorter and appealing to every student. (T178)

The Skill-based theme does not consist of any subthemes. In the Skill-based theme, teachers identified the new generation item concept with expressions such as "Stem, Steam, Outcome-based, Multidisciplinary." It is thought that teaching science, technology, engineering, art, and mathematics with an interdisciplinary approach, especially in the recent period, has also affected how teachers ask questions. Some of the teacher opinions in this theme are given below: "...It means a skill-based item that includes steam skills and is prepared with a multidisciplinary approach." (T45), "I had heard that it was different from the current education system, more exam-oriented." (T104)

The Taxonomic view theme consists of Bloom's taxonomy (13), higher level skills (12), 21st century skills (3) and learning outcome based (4) subthemes. In Bloom's taxonomy subtheme, teachers evaluated the new generation items in terms of their levels. Accordingly, they stated that this concept is at the comprehension, application, synthesis, and evaluation level. Some of the teacher opinions in this subtheme are given below:

They are items above the comprehension level. They are items involving measurement at the level of analysis and synthesis. (T4)

Item types that are at the levels of analysis and synthesis and even evaluation, rather than knowledge and comprehension, are very difficult for students, especially in the field of mathematics. (T60)

In the High-level skills subtheme, teachers defined the new generation item concept with expressions such as skill, thinking, mind-straining, making inferences, and understanding algorithms. Some of the teacher opinions in this subtheme are given below:

To measure the application and high-level skills of the student other than measuring memorization knowledge. (T17)

It is an item that requires high-level thinking skills. It is a problem situation that requires not knowledge but how knowledge should be used. (T28)

In the 21st century skills subtheme, teachers associated this concept with the skills necessary for the modern age. Some of the teachers' views on this subtheme are given below:

It means ensuring that 21st century skills (such as scientific literacy, critical thinking, creativity) are acquired. (T43)

New generation items are those that address higher skill levels based on problem solving, which is one of the 21st century skills required by the millennium age. (T188)

In the Learning outcome based subtheme, they tried to explain the new generation items with explanations such as being related to multiple learning outcomes, outcome-item interaction, using more than one piece of knowledge, and discriminating between those who know and those who do not know. Some of the teacher opinions on this subtheme are given below: "I think of the item that distinguishes the knower from the learner and tests many learning outcomes." (T54), "Problem solving method using more than one knowledge and skills..." (T140)

The Reading comprehension-based theme does not consist of any subthemes. In this theme, teachers tried to define the new generation item concept with statements such as reading comprehension, rapid reading, visual reading, transferring what they read, interpreting what they read, and reading skills. Some of the teacher responses to this theme are given below:

These are items that require the student to move away from the classical memorization method and use his/her ability to understand and interpret what he/she reads and to combine this skill with his/her knowledge. (T39)

Reading means understanding, interpreting, inferring. It includes longer sentences and visuals. (T44)

The theme of Restructuring knowledge consists of the subthemes of Use of knowledge (13), interpretation of knowledge (15), and thinking differently (5). In the subtheme of Use of knowledge, teachers defined the concept of new generation items as establishing relationships, different from rote memorization, creating connections, evaluating knowledge, analyzing knowledge and what is given to solve the problem. Some of the teachers' views on this subtheme are as follows: "New generation item means better use of knowledge." (T10), "It measures the ability to use data by giving prior knowledge, the ability to make interpretations..." (T95)

In the subtheme of Interpretation of knowledge, teachers tried to explain the new generation item concept with statements such as interpreting, explaining, inquiring, not asking direct knowledge, evaluating knowledge, and analyzing knowledge. In this subtheme, some of the teachers expressed the following opinions:

The new generation item is the student answering the item by interpreting it. It is to move away from the memorization approach. It is to make them understand. (T86)

In new generation items, the student does not need to memorize knowledge; instead, he/she answers the items by analyzing the knowledge he/she has learned for years and adding his/her interpretation. (T149)

In the subtheme that makes you Thinking differently, statements such as reasoning, deduction, comprehension and progressive thinking were used to explain the concept of new generation items. Some teachers' opinions on these expressions in this subtheme are given below: "In addition to knowledge, logic, reasoning, intelligence, comprehension of interdisciplinary collaboration." (T69), "The type of item that measures students' reasoning skills..." (T137)

The Daily life skills-based theme consists of Real world and daily life (10) and Gain-oriented (6) subthemes. In the actual world and daily life subtheme, the teachers tried to explain the concept of new generation items as experiencing by doing, real life, outdoor life, and daily life skills. Some of the teachers' views on this subtheme are given below: "The items that measure students' conceptual understanding and their ability to relate knowledge to daily life..." (T21), "...an item that makes connections with daily life..." (T82)

In the Gain-oriented subtheme, statements such as multidisciplinary, multilearning outcomes, and multibehaviours were related to the concept of new generation items. Teacher opinions on this subtheme are as follows: "... combining multiple learning outcomes..." (T52), "It includes learning outcomes from more than one discipline and contains interdisciplinary knowledge..." (T107)

The New view theme does not contain a subtheme. When the teacher responses under this theme were examined, it was seen that concepts such as adapting to the times, artificial intelligence assisted, and next generation humans were used. Some of the opinions of teachers on this theme are given below: "The items that adapt to the time..." (T22), "The item that chooses the person sought after in the new

generation...” (T155), “...the modern item that selects brains that create ideas that do that artificial intelligence cannot do. (T155)”

The theme of Reading comprehension-based," and the subtheme of "comprehension" in Kablan and Bozkuş's (2021) research, the theme of "reading comprehension skills" in Çaldıran and Özkan's (2022) research, and the subtheme of "reading and comprehension" in Kaya and Şahin's (2023) research are highly similar to the responses given in this research.

Discussion, Conclusion, and Suggestions

This research analyzed teachers' views on new generation items according to various variables. While 117 out of 192 teachers pointed out the existence of a new generation item concept, 92 of them stated that there was no such concept. It was observed that there was no significant difference between the teachers' responses regarding the concept of new generation item and their branches, occupational period, and whether they attended any course/seminar/workshop. A situation similar to this result is not found in the literature because there is no research examining the relationship between the new generation item concept and these variables.

In the qualitative section, where teachers' views on the concept of new generation items were taken, seven themes were identified. According to these themes, new generation items are mainly formal in measuring reading comprehension and higher-order thinking skills. According to teachers, new generation items consist of complicated structures of long texts, visuals, and various scenarios in which reading comprehension predominates, measuring higher-order thinking skills and incorporating many disciplines. In Erden's (2020) research, teachers' association of new generation items with learning outcomes is similar to the statements in the learning outcome-based subtheme in the taxonomic view theme of this research. In Çaldıran and Özkan's (2022) research, statements such as difficulty and complexity mentioned in student opinions are similar to the answers given in the formal perspective theme of the study. At the same time, when the old-type items were compared with the new items, it was seen that the students used expressions similar to the teachers' in this research, such as length, difficulty, and high-level thinking. In Kaya and Şahin's (2023) research, while Turkish teachers stated that the new generation items were incompatible with the learning outcomes of the curriculum, they used the statements of over-achievement or high-level thinking skills. These statements are similar to the statements in the taxonomic perspective theme of this research. The subthemes of problem-solving, critical thinking, and interpretation under the theme of cognitive skills in Şahan and Şahin's (2023) research are similar to the statements given in the higher-order skills subtheme under the taxonomic view theme and the reading comprehension-based theme in this research. The statements in the proximity to the life subtheme of the contextual features theme of the same research are similar to the statements in the actual world and daily life subtheme of the daily life skills-based theme of this research.

According to the opinions of teachers and the studies conducted, the concept of new generation items corresponds to the concepts of skill-based items used by MoNE, context-based items with studies in the literature, and item-like items used in PISA. Erden (2020), Kablan and Bozkuş (2021), Karakeçe (2021), Kertil et al. (2021), Kılcan (2021), Sanca et al. (2021), Çaldıran and Özkan (2022), İlkörücü and Altaş (2022), Karabulut, Tosunbayraktar and Kariper (2022), Şan and İlhan (2022), Yiğit, Deveci and Dadandı (2022), Türkel (2022), Ceylan and Orhan (2023), Şahan and Şahin (2023) stated that such items are called new generation items in their studies. However, the concept of new generation items is not found in any of MoNE's (2018, 2023) documents on measurement and evaluation processes. In this process that emerged with the MoNE (2018), 2023 Vision Document, this concept has never been used in the competency-based assessment and evaluation process. In addition, the concept of new-generation items is not mentioned in the formative assessment and evaluation process specified by MoNE (2023) in the new assessment and evaluation legislation. When this concept was searched in Turkish and English in international refereed journals published in assessment and evaluation indexed in indexes such as WoS and Scopus, no research was found. In addition, when the studies of Ahmed and Pollitt (2007), Chi, Wang, and Liu (2022), Hanberger (2014), and Roehl (2015), which are stated to be related to new generation items in the international literature, it is seen that the concept of "new generation item" is not included in their content in any way. When these studies were examined meticulously, it was determined that their content was context-based or based on reviewing

the items used in PISA. Accordingly, the concept of a new generation item is not used in local or international literature in a way identical to the concepts of competence, skill, or context-based or formative assessment and evaluation, which should be seen as evidence that this concept is not scientific.

An essential limitation of scientific research in the social sciences is the accurate construction and discussion of concepts. Concepts cannot be formed according to a scientist's preferences. There are specific criteria to be considered in this regard. Concepts are developed and named by these criteria, and no scientist can define that concept with a different name (Yalçın, 2017). Accordingly, to accept the existence of the new generation item, it needs to be rearranged according to the concept formation criteria.

The results obtained from the quantitative data and the results obtained from the qualitative data of the study were analyzed together, and it was seen that there was no significant relationship between the related variables. There was no significant difference between the fields of the teachers who stated that the new generation item concept existed, their occupational period, and whether they had taken any measurement and evaluation course or not. Accordingly, the variables examined do not affect the presence or absence of this concept.

Accordingly, it is recommended that item writers, teachers, and publishing organizations should not use the new generation item concept, which the researchers of this study claim is not a scientific concept. In addition, it is suggested that other researchers should know the historical process and basic knowledge of the relevant field well when generating concepts in a scientific field and should correctly determine the cognitive and intellectual meanings of the concepts.

Acknowledgment

Copyrights: The works published in the e-Kafkas Journal of Educational Research are licensed under a Creative Commons Attribution-Non-commercial 4.0 International License.

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 all ethical violations.

Author Contributions: Conceptualization, Eray Selçuk; methodology, Eray Selçuk; validation, Eray Selçuk, Ferhat Karakaya and Mehmet Yılmaz; analysis, Eray Selçuk; coding themes, Eray Selçuk, Ferhat Karakaya and Mehmet Yılmaz; writing, review and editing, Eray Selçuk and Ferhat Karakaya; supervision, Mehmet Yılmaz; project administration, Mehmet Yılmaz.

Funding: This research received no funding.

Institutional Review Board Statement: Ethical approval was obtained from Yozgat Bozok University Social and Human Sciences Ethics Committee to conduct the research with the decision dated 20/04/2022 and no. 32/40.

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

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

References

- Agresti, A. (2018). *An Introduction to Categorical Data Analysis*. John Wiley & Sons.
- Ahmed, A. & Pollitt, A. (2007). Improving the quality of contextualized questions: An experimental investigation of focus. *Assessment in Education Principles Policy and Practice*, 14(2), 201–232. doi:10.1080/09695940701478909
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Prentice Hall/Pearson Education.
- Brookhart, S. M. (2010). *Assess higher-order thinking skills in your classroom*. Alexandria, VA: ASCD.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. A., Karadeniz, Ş. & Demirel, F. (2013). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Akademi Yayıncılık.
- Ceylan, G., & Orhan, A. T. (2023). Beceri temelli fen sorularına yönelik fen bilimleri öğretmenlerinin görüşlerinin incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 19(1), 84-11. doi:10.17860/mersinefd.1193968
- Chi, S., Wang, Z. & Liu, X. (2022) Assessment of Context-Based Chemistry Problem-Solving Skills: Test Design and Results from Ninth-Grade Students. *Research in Science Education*, 53, 295-318. doi:10.1007/s11165-022-10056-8
- Creswell, J. W. (2020). *Nitel araştırma yöntemleri beş yaklaşıma göre nitel araştırma ve araştırma desenib*. Ankara: Siyasal Kitabevi.
- Creswell, J. W. (2021) *Karma Yöntem Araştırmalarına Giriş* (Çeviri Ed. M. Sözbilir). Ankara: Pegem Akademi Yayıncılık.
- Çaldıran, Ü. E., & Özkan, M. (2022). Çevrimiçi öğretim ortamlarında yeni nesil sorulara dayalı fen öğretimine ilişkin öğrenci görüşlerinin belirlenmesi. *Türk Akademik Yayınlar Dergisi (TAY Journal)*, 6(2), 311-328. doi:10.29329/tayjournal.2022.510.07
- Elmas, R., & Eryılmaz, A. (2015). Bağlam temelli fen soru yazımı: Kriterler ve efsaneler. *Kuramsal Eğitimbilim Dergisi*, 8(4), 564-580. doi:10.5578/keg.10135
- Erden, B. (2020). Türkçe, matematik ve fen bilimleri dersi beceri temelli sorularına ilişkin öğretmen görüşleri. *Academia Eğitim Araştırmaları Dergisi*, 5(2), 270-292.
- Gelbal, S. (1994). pMadde güçlük indeksi ile Rasch modelinin b parametresi ve bunlara dayalı yetenek ölçüleri üzerine bir karşılaştırma. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 10(1994), 85-94.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Lawrence Erlbaum Associates, Inc.
- Hanberger, A. (2014). What PISA intends to and can possibly achieve: A critical programme theory analysis. *European Educational Research Journal*, 13(2), 167–180. doi:10.2304/eej.2014.13.2.167
- İlkörücü, Ş., & Altaş, R. (2022). Evaluation of 8th-grade students' approaches to solving new-generation science questions based on variable determination from scientific process skills. *Online Science Education Journal*, 7(1), 17-27.
- Kablan, Z., & Bozkuş, F. (2021). Liselere giriş sınavı matematik problemlerine ilişkin öğretmen ve öğrenci görüşleri. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 17(1), 211-231. doi:10.17860/mersinefd.800738
- Kabuklu, Ü. N., Yüzbaşıoğlu, M. K., & Kurnaz, M. A. (Nisan, 2019). *Fen eğitimiyle alakalı araştırmalarda bağlam temelli soru yazma ölçütlerinin belirlenmesi*. Uluslararası Fen, Matematik, Girişimcilik ve Teknoloji Eğitimi Kongresi.
- Karabulut, H., Tosunbayraktar, G., & Kariper, İ.A. (2022). Ortaokul öğrencilerinin beceri temelli (yeni nesil) fen bilimleri sorularına yönelik görüşlerinin incelenmesi. *EDUCATIONE*, 1(2), 301-320.
- Karakeçe, B. (2021). *Ortaokul matematik öğretmenlerinin beceri temelli sorulara ilişkin değerlendirmeleri*. [Yayınlanmamış yüksek lisans tezi]. Gaziantep Üniversitesi, Gaziantep, Türkiye.
- Kaya, A., & Şahin, Ç. (2023). Sınıf öğretmenlerinin yeni nesil Türkçe sorularına ilişkin görüşleri. *Ana Dili Eğitimi Dergisi*, 11(2), 437-453. doi:10.16916/aded.1236839
- Kertil, M., Gülbagcı Dede, H., & Ulusoy, E. G. (2021). Skill-based mathematics questions: What do middle school mathematics teachers think about and how do they implement them? *Turkish*

- Journal of Computer and Mathematics Education*, 12(1), 151-186.
doi:10.16949/turkbilmat.774651
- Kılcan, T. (2021). Yeni nesil matematik sorularına ilişkin tutum ölçeği geliştirme: Geçerlik ve güvenilirlik çalışması. *Anadolu Kültürel Araştırmalar Dergisi*, 5(2), 170-180
- Kirk, R. E. (2013). *Experimental Design: Procedures for the Behavioral Sciences*. Sage Publications.
- Koch, G. G. (1982). *Intraclass correlation coefficient*. In Encyclopedia of Statistical Sciences (Vol. 4, pp. 213-217). Wiley.
- Nasırlıel, E. (2020). *8. sınıf öğrencilerinin bağlam temelli basınç sorularını çözme süreçleri*. [Yayımlanmamış yüksek lisans tezi]. Tokat Gaziosmanpaşa Üniversitesi, Tokat, Türkiye.
- McMillan, J. H. (2014). *Classroom assessment: principles and practice for effective standard-based instruction*. New York, NY: Pearson.
- MEB [MoNE] (2018). *2023 Eğitim Vizyonu*. Erişim adresi: https://www.gmka.gov.tr/dokumanlar/yayinlar/2023_Egitim%20Vizyonu.pdf Erişim tarihi:17.05.2023
- MEB [MoNE] (2023). Milli Eğitim Bakanlığı Ölçme ve Değerlendirme Yönetmeliği. *T.C. Resmî Gazete* (32304, 9 Eylül 2023). Erişim adresi: <https://www.resmigazete.gov.tr/eskiler/2023/09/20230909-2.htm>
- Merriam, S. B (2013). *Nitel araştırma: Desen ve uygulama için bir rehber* (S. Turan, Çev. Edit.). Ankara: Nobel Akademik Yayıncılık.
- Miller, D., Linn, R., & Gronlund, N. (2009). *Measurement and Assessment in Teaching*. New Jersey: Pearson.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded Sourcebook*. (2nd ed). Thousand Oaks, CA: Sage.
- Osterlind, S. J. (1990). Toward a uniform definition of a test item. *Educational Research Quarterly*, 14(4), 2-5.
- Osterlind, S.J. (2002). What is constructing test items? In S. J. Osterlind (Ed.), *Constructing test items: Multiple-choice, constructed-response, performance, and other formats* (pp. 1–16), Springer. doi:10.1007/0-306-47535-9_1
- Pallant, J. (2015). *SPSS Survival Manual: A Step by Step Guide to Data Analysis using SPSS*. (4th ed.). Maidenhead: Open University Press.
- Roehl, T. (2015) What PISA measures: some remarks on standardized assessment and science education. *Cultural Studies of Science Education* 10, 1215–1222. doi:10.1007/s11422-015-9662-z
- Sak, M. (2018). *Ortaokul öğrencilerinin ışık konusundaki bağlam temelli sorular ile geleneksel soruları cevaplama düzeylerinin karşılaştırılması*. [Yayımlanmamış yüksek lisans tezi]. Kocaeli Üniversitesi, Kocaeli, Türkiye.
- Sanca, M., Artun, H., Bakırcı, H. & Okur, M. (2021). Ortaokul beceri temelli soruların yeniden yapılandırılmış Bloom Taksonomisine göre değerlendirilmesi. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 219-248.
- Selçuk, E. (2022). Eray Selçuk ile Beceri Temelli Çoktan Seçmeli Maddelere Geçiş Süreci. Youtube: <https://www.youtube.com/watch?v=Dat4TzIIZyc> Erişim Tarihi: 17.05.2023.
- Şahan, H., & Şahin, Ç. (2023). İlkokul 4. sınıf matematik dersindeki yeni nesil sorularının öğretmen görüşlerine göre incelenmesi. *Uluslararası Bilim ve Eğitim Dergisi*, 6(1), 83-98. doi:10.47477/ubed.1231419
- Şan, S., & İlhan, N. (2022). Fen bilimleri dersi beceri temelli sorulara (yeni nesil) yönelik kuramsal ve kavramsal çerçeve. *İnönü Üniversitesi Eğitim Bilimleri Enstitüsü Dergisi*, 9(17), 17-36. doi:10.29129/inujse.1089655
- Taşdemir, M. (2015). Öğretmen adaylarının Türk millî eğitim sistemi üzerine algıladıkları sorunlar. *Turkish Studies-International Periodical for the Languages, Literature and History of Turkish or Turkic*, 10(7), 881-898. doi:10.7827/TurkishStudies.7898
- Tekbıyık, A., & Akdeniz, A. R. (2010). Bağlam temelli ve geleneksel fizik problemlerinin karşılaştırılması üzerine bir inceleme. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 4(1), 123-140.
- Thorndike, L. R. (1982). *Applied Psychometrics*. Boston: Houghton Mifflin Co.

- Türkel, R. (2022). İslam mezhepleri tarihi ders kazanımlarının ölçme ve değerlendirmesinde beceri temelli çoktan seçmeli madde (yeni nesil soru) kullanımı. *Mezhep Araştırmaları Dergisi*, 15(2), 755-768. doi:10.18403/emakalat.1223301
- Ural, A. (2016). Yarışmacı eğitim anlayışının etkileri üzerine bir çözümleme. *Journal of Critical Pedagogy*, 43, 19-24.
- Ünal, M. (2019). *PISA sınavlarının özelliklerinin fen bilimleri öğretmenlerinin hazırlamış oldukları sınav soruları ile karşılaştırılması: PISA kültürünü yaygınlaştırma model önerisi*. [Yayımlanmamış doktora tezi]. Uludağ Üniversitesi, Bursa, Türkiye.
- Yalçın, A. (2017). Bilimsel çalışmalarda kavram oluşturma ve sözsüz iletişim kavramları üzerine. *Avrasya Terim Dergisi*, 5(2): 74-79
- Yıldırım A., & Şimşek H. (2021). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri* (12. Baskı). Ankara: Seçkin Yayıncılık.
- Yiğit, N., Deveci, İ. & Dadandı, N. (2022). Yeni nesil fen bilimleri sorularına yönelik algı ölçeğinin geliştirilmesi. *Van YYÜ Eğitim Fakültesi Dergisi*, Şubat Özel Sayı, 108-130. doi:10.33711/yyuefd.1068089