Original Research

SAGE Open
April-June 2026: 1–23

DOI: 10.1177/21582440261431848
journals.sagepub.com/home/sgo
Sage

Assessment of High School Students' Formulating Skills: A Framework Proposal

Hasan Temel[1]

Abstract

Formulating is a crucial skill for transforming real-life situations into mathematical representations and reconnecting mathematical abstractions to real contexts. This skill plays an essential role in mathematical literacy, problem-solving, modelling, and STEM practices. This study aims to examine high school students' formulating skills and to propose a framework for evaluating them. The Formulating Test developed for this research was administered to 561 students to conduct validity and reliability analyses. After establishing the test's measurement properties, the formulating skills of an additional 431 students from different high school types were analysed in relation to the proposed framework, school type, grade level, and gender. The findings indicate that students' formulating skills are generally below the expected level, consistent with previous research. Moreover, results revealed that (1) students performed better in the mathematical sense-making dimension than in other dimensions, (2) school type had a significant effect on formulating skills, (3) grade level had a small effect, and (4) gender differences were not significant. These findings highlight the need for more comprehensive studies to support the development of formulating skills. Overall, this research provides an important foundation for strengthening formulating skills in future studies and educational practices.

Plain Language Summary

What is the Current Level of Formulating Skills Among High School Pupils?

This study examines high school students' ability to transform real-life situations into a mathematical structure and to adapt mathematical generalisations to real life, that is, their ability to formulate. Formulating involves understanding a situation, selecting the necessary information, and transforming it into a mathematical problem. To assess this skill, a special 15-question test was developed using PISA problems related to the formulating process within the research framework. The test was first adapted for a group of students, then administered to 431 students from different high schools, and the results were analysed. The findings indicate that students are successful at translating given situations into mathematical problems; however, they struggle to generate new real-life problems using mathematical thinking. This situation reveals that students need more support in constructing their own problems and structuring their mathematical thinking. These results indicate that students' ability to formulate real-life problems and structure their mathematical thinking needs further development. Therefore, it is recommended that more emphasis be placed on modelling activities, STEM applications, and open-ended problem-setting exercises to support skill development.

[1]Balıkesir University, Türkiye

Corresponding Author:
Hasan Temel, Department of Mathematics Education, Necatibey Faculty of Education, Balıkesir University, Dinkçiler, Soma Street, Balıkesir 10145, Türkiye.
Email: hasan.temel@balikesir.edu.tr

Data Availability Statement included at the end of the article

Keywords
mathematics education, formulating skills, high school students, mathematical literacy, secondary education, problem formulation

Introduction

Today, mathematics is seen as a prerequisite course and a fundamental aspect of literacy for being a stakeholder in a modern society (National Council of Teachers of Mathematics [NCTM], 1989). It is now necessary to move away from memorising procedures in mathematics and focus on mathematical thinking, reasoning, and problem-solving (Cai & Howson, 2012). With this evolving perspective, mathematics is not merely a collection of abstract operations but a fundamental literacy area that individuals need throughout their lives. In other words, mathematics is a tool that provides individuals with basic skills such as analytical thinking, problem-solving, and reasoning, which STEM education and critical thinking focus on, rather than a discipline that individuals only see as abstract and logical (Maass et al., 2019).

Mathematics plays a crucial role in many fields, including science, technology, engineering, architecture, robotics, coding, art, agriculture, and medicine. It is effectively used for everyday analysis, prediction, inference, and interpretation. The use of mathematics is not limited to quantitative fields such as science, technology, engineering, and medicine, but also plays an active role in the social sciences. In the social sciences, mathematical modelling and analysis are utilised in numerous contexts, ranging from human behaviour to societal concerns, and from economic expectations to political processes (Coleman, 1990; Hedström, 2005). Mathematics provides a scientific basis for drawing conclusions, formulating policy, enhancing the quality of research processes, and making predictions in the social sciences. In this context, mathematics is at the core of communication, reasoning, correlation, and problem-solving in the social sciences as much as in the natural sciences. In this regard, cognitive processes such as decision-making, problem-solving, reasoning, and interpretation in daily life heavily rely on mathematical thinking. Developing these skills is essential for an individual's success in both social and academic settings (Niss & Højgaard, 2019). Therefore, contemporary mathematics education aims not only to enhance students' computational abilities but also to equip them with the skills needed to analyse real-life situations, convert these situations into mathematical models, and reinterpret the results. In this framework, mathematics curricula focus on developing individuals' critical and creative thinking, reasoning, problem-formulating, and solution-generating skills, enabling them to become mathematically literate (Leung et al., 2014; Tesfamicael & Enge, 2024).

In recent years, the concept of mathematical literacy has received considerable attention, particularly through the Organisation for Economic Co-operation and Development's (OECD) Programme for International Student Assessment (PISA). PISA evaluates students' abilities to formulating, employing, and interpreting mathematical processes, emphasizing how effectively they can translate real-life problems into mathematical frameworks (OECD, 2023). The ability to formulating is the foundation of mathematical thinking, a critical problem-solving strategy, and an essential component of the modelling process. Research shows that students often struggle specifically in the formulating stage (Tasarib et al., 2025). These findings highlight the need for a more comprehensive approach to teaching the skill of formulating in educational programs.

Background

Formulating is a fundamental thinking skill that allows individuals to structure abstract concepts and develop logical reasoning processes for problem-solving (Mason et al., 2010). The ability to generalise and abstract relationships between events and objects forms the foundation of mathematical thinking. However, converting these relationships into systematic mathematical expressions requires specific formulating skills (Tall, 2013).

For instance, recognising the Fibonacci sequence in nature demonstrates mathematical thinking, while translating it into a symbolic expression requires formulation. In the problem-solving process, formulating is essential for both developing strategies and generating solutions (Polya, 1945; Schoenfeld, 1985). Consequently, formulating plays a crucial role in helping students understand the problems they face, develop paths to solutions, and construct mathematical structures (Luna & Díaz, 2023).

Formulating skill is at the core of mathematics education. It is considered an important component across many approaches, such as problem-solving, mathematical modelling, realistic mathematics education, STEM education, and mathematical literacy (Figure 1). The problem-solving literature emphasises that the process involves formulating the problem and generating a solution (Santos-Trigo, 2024), which aligns with the stages proposed by Polya and Schoenfeld. In mathematical modelling studies, it has also been noted that the formulating process plays a decisive role in modelling real-life situations (Berry & Houston, 1995; Stillman et al., 2007). Similarly, the mathematical literacy framework

Figure 1. Mathematical processes involving the formulating.

emphasises individuals' ability to express, reason, and solve daily-life problems mathematically; it states that formulating skills are at the core of this transformation (MoNE, 2023; OECD, 2013, 2023).

Polya (1945) and Schoenfeld's (1985) framework for problem solving provides a starting point for understanding the place of formulating in mental processes, but it does not include cognitive aspects such as the formation of the problem situation, how this situation corresponds to the mathematical world, how problems are constructed, and a detailed explanation of the formulating process. Similarly, while Krutetskii's (1976) analytical descriptions of mathematical abilities are valuable, this model is also limited in explaining the processes of formulating presented in the context of today's realistic mathematics education, modelling, STEM, and mathematical literacy. Therefore, this study aims to reposition the skill of formulating within the framework of contemporary research, based on classical theoretical approaches.

In the context of realistic mathematics education, formulating skills are used as an effective tool in horizontal and vertical mathematisation processes (Artut & Bal, 2016; Zulkardi, 2000). In the context of STEM education, formulating is a critical stage in integrating science, technology, engineering and mathematics disciplines, defining scientific research questions, and identifying engineering design problems (Kennedy & Odell, 2014). This process enables the transfer of problems in real-life contexts to the mathematical plane and the development of solutions by integrating interdisciplinary content. Similarly, in the 'computational thinking' approach, formulating the problem is defined as the first step in generating solutions through a computer (Wing, 2006); in this context, studies conducted within the framework of computational thinking-based mathematical problem-solving activities reveal how necessary the formulating skill is even at the primary education level (Ng & Cui, 2021).

Additionally, research indicates that students' ability to solve mathematical problems and to engage in mathematical modelling and problem posing largely depends on their formulating skills (Leikin et al., 2025; Santos et al., 2024). Several cognitive challenges have been identified, particularly in understanding the problem context, recognising relationships, and creating appropriate mathematical expressions (Aguirre et al., 2024). Consequently, formulating is closely linked to the development of higher-order thinking skills, such as reasoning, generalisation, and abstraction (Niss & Højgaard, 2019).

Formulating and Proposing a Framework

Mathematical formulating is a high-level cognitive process that enables individuals to transform contextual situations into mathematical structures logically and symbolically. This process relies on fundamental cognitive skills such as abstraction, establishing relationships, recognising structures, and generalisation. Krutetskii (1976) states that formulating is based on students' ability to recognise structures and relationships within content, while the OECD (2013) defines formulating as an individual's ability to structure a contextual situation mathematically, thereby explaining its position in mathematical literacy. Within these perspectives, formulating is seen as a fundamental theoretical structure at the core of mathematical thinking.

Formulating skills include (a) Mathematical sense-making: making mathematical sense of situations given in a real-world context (OECD, 2023), (b) Structuring and generalising: reaching generalisations in the process of transforming mathematical expressions into a systematic structure (Tall, 2013), (c) Interpretation of Mathematical Abstractions: interpretation of abstract concepts in mathematics to real life world contexts (Blum & Niss, 1991), and (d) Real-life problem creation: transforming the applications of mathematical abstractions in real life into problem situations (Figure 2).

(a) *Mathematical sense-making:* This stage, which is a component of formulating, requires the conversion of a problem situation given in a real-life

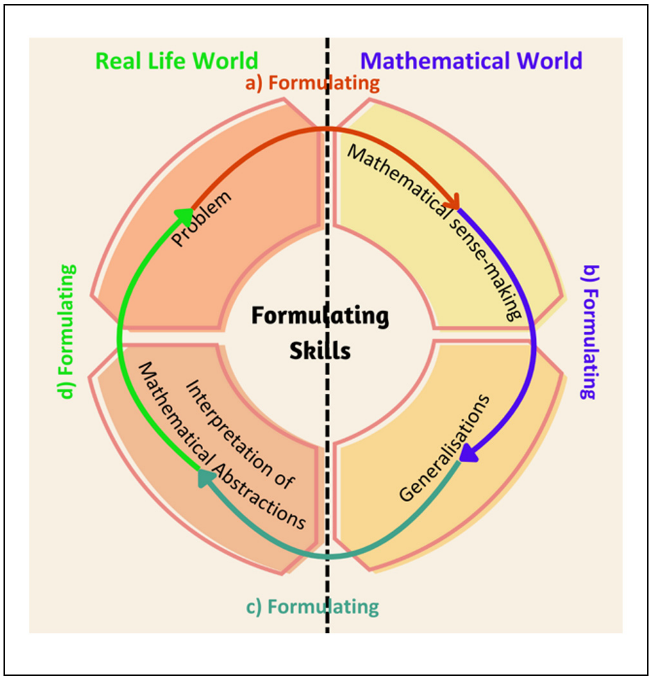

Figure 2. Formulating process model.

context into mathematical expressions and the representation of the problem in the context of images, diagrams, and symbolic expressions to achieve a mathematical appearance. The mathematical sense-making process corresponds precisely to the ability to formulate in mathematical literacy (OECD, 2023). This process can be associated with horizontal mathematisation in realistic mathematics education and mathematical modelling. The horizontal mathematisation process refers to the transition from real-life contexts to the world of symbols (Doorman & Gravemeijer, 2009). In other words, horizontal mathematisation is the process of transforming concrete problems into mathematical concepts (Treffers, 2012). Similarly, organising and reorganising mathematical knowledge to understand and analyse real-life situations is associated with modelling (Ferri, 2006). The modelling process involves creating a physical, symbolic, or abstract model of a situation (Lesh & Doerr, 2003). For these transformation processes to be carried out successfully, the ability to formulate (the mathematical sense-making process) must be applied. For example, when a student is given the problem 'there are 15 eggs in each box,' the student's ability to model this as '15 × ' and convert it into a mathematical form demonstrates that this process has been carried out correctly. In this context, converting informal information presented in a real-life setting into a mathematical structure requires a formulating process (mathematical sense-making). The mathematical sense-making process emerges as the stage that enables understanding and solving problem situations within a mathematical structure and contributes to the development of mathematical thinking skills (Altun, 2015).

(b) *Structuring and generalising:* The structuring and generalising process involves students analysing problem situations derived from real-life contexts, transforming these situations into mathematical structures, identifying patterns and arriving at generalisations from these patterns (Tall, 2013; Treffers, 2012). This process is particularly associated with vertical mathematisation and enables students to transition from concrete contexts to abstract mathematical representations. Mathematical modelling is a fundamental tool that supports this transition and facilitates the development of generalisation and abstraction skills (Blum & Leiss, 2007). Students' development of general formulas based on specific numerical patterns—for example, arriving at the formula $n(n + 1)/2$ for the sum of numbers from 1 to n—demonstrates how the process of formulating works in the context of structuring and generalising. This process plays a critical role in enabling students to not only solve problems but also to structure and generalise mathematical concepts (Altun, 2020; Köroğlu & Yeşildere, 2004). Therefore, the ability to reach generalisations through formulating in mathematics education is pedagogically important in terms of deepening conceptual understanding and developing higher-order thinking skills.

(c) *Interpretation of mathematical abstractions:* This process involves relating mathematical generalisations and abstractions to real-life situations in a formulaic context. The process of formulating involves relating mathematical generalisations and abstractions to real-life situations. Applications developed in real-life contexts require the effective use of this process to help students understand abstract mathematical concepts (Blum & Niss, 1991). The mathematics education literature emphasises that linking abstract concepts to concrete experiences strengthens conceptual understanding (NCTM, 2000; Van Den Heuvel-Panhuizen, 2003). In this context, the formulating process enhances the durability of learning by enabling students to construct abstract mathematical structures through

Table 1. Summary of the Formulating Process Framework.

Dimension	Description	Measurability
(a) Mathematical sense-making	The process of converting real-life situations into mathematical structures	Production of mathematical equations, graphs, tables, figures, etc.
(b) Structuring and generalising	The process of making mathematical structures abstract/generalisable	Production of generation structures and formulas
(c) Interpretation of mathematical abstractions	The process of relating mathematical generalisations and abstractions to real life	Ability to produce applications for the use of mathematical generalisations and abstractions in real life
(d) Real-life problem creation	The process of generating mathematical problems from real life	Production of generating mathematical problems in the context of formulating a situation in real life.

concrete problem situations. Let us consider the formula $n(n + 1)/2$ for the sum of consecutive numbers as an example of the interpretation of mathematical abstractions. Presenting this mathematical generalisation directly may remain abstract for students. However, when this formulating process is presented in the context of a problem set up in a schoolyard where one person sits on each step in sequence, the formulating process can be triggered by asking students to find the total number of people. In this way, students can understand an abstract mathematical concept by constructing it through a contextual problem and arrive at a generalisation. Such contextual approaches not only develop formulating skills but also enhance the depth and durability of learning.

(d) *Real-life problem creation:* This process involves formulating real-life situations as problems. Formulating problems from real life to a real-life context requires a formulating process. Because the real-life problem creation process involves analysing real-life situations and expressing them as problems, it can also be associated with the first stage of mathematical modelling. Mathematical modelling starts with analysing real-life situations (Blum & Leiss, 2007). The context in which the situations will be handled is revealed at this stage. In its definition of problem-solving, the NCTM (1989) supports the view that mathematics can be used as a tool to formulate events within and outside mathematics and that creating problems in real-life contexts requires a formulating process. If we explain this process with an example, the issue of reducing energy consumption can be addressed through an investigation within the scope of environmental sustainability. In this context, a problem can be formulated to create a problem context that reduces energy consumption. While expressing this problem, the situation can be presented in the context of a formula for carbon emission, which represents the environmental impact of energy consumption. Problems can arise when it is revealed that carbon emissions increase linearly with energy consumption. The problem formulation process requires a formulating process (Table 1).

Rationale

NCTM (2000) emphasises that problem-solving, reasoning, relating, and communicating should be central to mathematics education. According to the Ministry of National Education (MoNE, 2018), students' understanding of mathematics in real-life contexts is essential for problem-solving and higher-order thinking skills. These perspectives highlight that today's mathematics education requires not only the development of procedural skills but also higher-order thinking skills that involve transforming real-life situations into mathematical structures, understanding abstract concepts, and developing new problem-solving scenarios. In this context, the ability to formulate is central to problem-solving processes, along with mathematical literacy (OECD, 2023), which is necessary for transforming real-life situations into mathematical structures, and modelling (Blum & Leiss, 2007), which helps concretise abstract concepts.

The OECD's (2023) PISA 2022 results show that students in Türkiye experience severe difficulties in mathematical literacy, particularly in formulating, and perform below the average of participating countries. This situation also reveals that teaching practices expected by the curriculum, such as mathematical reasoning, generalisation, and modelling, have not yet been sufficiently developed in practice. Therefore, this issue requires immediate attention. Furthermore, the difficulties students face in translating real-life situations into mathematical structures point to a systemic instructional gap that requires urgent intervention.

The limited research on high school students' formulating skills highlights the absence of comprehensive scales and frameworks for assessing this ability. Therefore, examining high school students' formulating skills by focusing on the difficulties they encounter can shed light on current learning outcomes and open the door to designing curricula and teaching processes in this direction. While the literature has identified a lack of scales and theoretical models aimed at improving formulating skills, especially at the high school level, a systematic framework to address this shortcoming has yet to be proposed. Many existing studies often examine isolated components of formulating within the context of mathematical literacy or primarily focus on middle school students. This approach leaves a conceptual and empirical gap regarding the development of formulating skills in older students. The existing literature on formulating has significant gaps, highlighting the urgent need for studies that conceptualise formulating as a multidimensional construct and provide a solid empirical basis for its assessment.

Related Literature and Aim of the Study

Mathematics teaching should provide students with opportunities to approach events from different perspectives, to reflect on their thinking, and to evaluate their procedures critically (Cai & Howson, 2012). Such reflections play a critical role, especially in the process of formulating new problems (Toh et al., 2023). However, studies reveal that students experience various difficulties in formulating problems (Levenberg & Shaham, 2014; Pérez et al., 2017). Ufer et al. (2009) stated that students have difficulty formulating mathematical proofs even at later stages of their schooling, while Pérez et al. (2017) reached similar results at the international level and emphasised that students are unable to formulate mathematical proofs independently. In addition, Ramdhani et al. (2019) stated that students struggle to express generalisations symbolically and develop rules based on patterns.

These difficulties in the formulating process are evident across different age and ability groups. Levenberg and Shaham (2014) found that approximately half of gifted students in grades 4–6 could not formulate verbal problems. Yang and Lin (2015) stated that Taiwanese students lack mathematical literacy, especially when formulating real-life problems, and cannot distinguish between relevant and irrelevant information. Edo et al. (2013), in their study focusing on high-level PISA problems, stated that students had difficulty identifying mathematical structures (patterns, relationships, orders) and formulating mathematical representations of these structures. Similarly, Dewantara et al. (2015) found that formulating was the process in which seventh-grade students made the most errors. Wildani (2020) reported that 47.3% of ninth-grade students' errors were related to the formulating process. These findings suggest that formulating is a central area of difficulty in mathematical literacy and problem-solving.

Most studies in the literature indicate that high school students experience particular difficulties in formulating. However, research on the proficiency of high school students in the formulating process is limited (Blum & Leiss, 2007). This study offers a unique contribution to mathematics education by proposing an empirically validated multidimensional framework for high school students' formulating skills. Unlike previous research, which generally treats formulating skills as a single-step cognitive process, this study conceptualises them through four distinct and interrelated dimensions. It operationalises these within a carefully developed assessment tool. The study not only proposes a new theoretical framework but also provides robust empirical evidence from a large, diverse sample, addressing the need for systematic tools and models in this under-researched area. In this regard, the primary objective of the research is to determine the levels of this skill among high school students and to examine how variables such as school type, grade level, and gender affect it.

Formulating is critical to mathematical achievement. Therefore, identifying the factors influencing this process can inform strategic planning to improve students' competencies in this area. In addition, considering the relationship between formulating and basic concepts such as mathematical literacy, problem solving and modelling, the study's findings are also expected to guide students' academic achievement and career development. In line with the data obtained, the aim is to review the curricula, restructure the teaching methods, and contribute to increasing the success of international assessment applications such as PISA.

Research Problem

The fundamental problem of this research: What is the level of high school students' formulating skills?

Within this scope, answers will be sought to the following sub-research questions (RQ):

RQ1. Is there a significant difference in the success levels of high school students across the sub-dimensions of the formulating process framework?

RQ2. Do high school students' formulating skills differ across school types?

RQ3. Do high school students' formulating skills differ by grade level?

RQ4. Do high school students' formulating skills differ by gender?

RQ5. Are there significant interaction effects of school type, grade level, and gender on students' formulating skills?

Hypotheses

H1: Students' performance across the four sub-dimensions of the proposed formulating framework significantly differs.

H2: High school students' formulating skills significantly differ according to school type.

H3: High school students' formulating skills significantly vary across grade levels, with higher grades demonstrating higher performance.

H4: There is no statistically significant difference in formulating skills between male and female students.

H5: School type, grade level, and gender have a statistically significant interaction effect on high school students' formulating skills.

Method

In this study, the survey method was used to determine the levels of high school students' formulating. Survey design is a non-experimental research approach that quantitatively describes individuals' attitudes, opinions and tendencies through a sample selected from a specific population (Creswell, 2013). The primary purpose of survey studies is to understand the population's characteristics and make statistical predictions accordingly (Fowler, 2014). This method, which is generally based on data collection tools such as questionnaires and interviews (Johnson & Christensen, 2012), can be applied cross-sectionally or longitudinally, depending on the time of data collection (Creswell, 2013). This study's cross-sectional survey design was preferred since the data were collected simultaneously.

As the research aimed to reveal the current state of students' formulating skills, a cross-sectional survey design was adopted using data collected in a single session. The research structure also incorporates comparative survey features, and comparisons were made of students' formulating performance by school type, class level, and gender. The research is theoretically structured as a descriptive study focusing on formulating, in line with the OECD's mathematical literacy framework.

Participants

This study was conducted with students from Anatolian High School, Science High School, Fine Arts High School, and Social Sciences High School in a city centre in the Marmara Region of Türkiye, with permissions obtained from the Ministry of National Education (MoNE). The provincial centre where the research was conducted has two Science High Schools, one Social Sciences High School, one Fine Arts High School, and over ten Anatolian High Schools. In the study, at least one school was selected from each type of school, and the applications were carried out in schools authorised for the research by the provincial directorate of national education. The study was conducted at the high school level because research on formulating skills generally focuses on the middle school level, and there are not enough studies on this age group. In addition, according to Piaget's cognitive development theory, high school students are in the abstract operations period, and it is accepted that high-level skills such as abstract thinking, problem-solving, and reasoning develop during this period. Since formulating is closely related to these skills, the high school level was determined as an appropriate target group. The fact that PISA is also aimed at the 15-year-old age group is another reason to conduct the study at this level.

There were 431 students from four different high school types: Science High School (108), Social Sciences High School (110), Fine Arts High School (108), and Anatolian High School (105). The 431 students participating in the study were selected from schools authorised to conduct the study under research permits issued by the Ministry of National Education. They comprised students from schools that voluntarily participated, providing the researcher with convenience in conducting the study. The selection of four different types of schools is based on student statistics cited in the National Education Statistics Formal Education 2023/2024 Report, published by MoNE (2024). According to the report, approximately 2 million of the approximately 2.16 million students enrolled in public high schools, excluding open education high school students, attend these four types of high schools. The selected school types cover the types of high schools attended by approximately 93% of students enrolled in secondary education institutions other than open education high schools. Another reason for selecting these school types with different academic orientations was to obtain comparable groups. The sample was selected using convenience sampling, with participants identified from easily accessible individuals (L. Cohen et al., 2005). In this context, the schools participating in the research were selected from institutions easily accessible to researchers. Students at these schools volunteered for the study. In addition, students who were not included in the scale development process were included in the study, and data collection was conducted independently of scale development.

Table 2. Demographic Information of the Participants.

High school type	Gender	Grade 9	Grade 10	Grade 11	Grade 12	Total
Science High School	Female	16	17	15	17	65
	Male	11	10	11	11	43
	Total	27	27	26	28	108
Social Sciences High School	Female	15	11	12	16	54
	Male	13	16	15	12	56
	Total	28	27	27	28	110
Fine Arts High School	Female	16	12	15	14	57
	Male	11	14	13	13	51
	Total	27	26	28	27	108
Anatolian High School	Female	11	10	12	17	50
	Male	14	16	14	11	55
	Total	25	26	26	28	105
Overall total	Female	58	50	54	64	226
	Male	49	56	53	47	205
	Total	107	106	107	111	431

To determine the study sample, statistical methods were used to ensure it was sufficiently representative of the target population, based on the MoNE's (2024) report, which indicates that approximately 2.16 million students are enrolled in high schools. Calculations were made assuming a 95% confidence level, a 5% margin of error, and an estimated p-value of .50, accounting for the maximum possible diversity in the population. The minimum sample size calculated using the method recommended by Taherdoost (2017) for survey studies and based on the Cochran formula is 385. This result is also consistent with the classic sample table of Krejcie and Morgan (1970). Based on this information, the sample size of this study, 431 students, is reasonably sufficient.

Additionally, a post-hoc power analysis was conducted to evaluate the statistical adequacy of the study's sample size of 431 participants (α = .05). The analyses indicate that the power is above 0.99 for medium effect sizes (Cohen's d = 0.50; Cohen's f = 0.25) for the t-test for independent samples and one-way analysis of variance (ANOVA). This result indicates that the study sample is more than sufficient to reliably detect the medium and large effects observed. It should also be noted that larger samples may be required to detect small effect sizes. The demographic information of the participants is presented in Table 2.

According to Table 2, of the 431 students who participated in the study, 226 were female (52.4%), and 205 were male (47.6%). When the distribution by gender is analysed in terms of school types, 60.2% of Science High School students are female and 39.8% are male; 49.1% of Social Sciences High School students are female and 50.9% are male; 52.8% of Fine Arts High School students are female and 47.2% are male; 47.6% of Anatolian High School students are female and 52.4% are male. These distributions show that a balanced sample was obtained with respect to gender overall.

When the participants were analysed by grade level, it was seen that a similar number of 9th, 10th, 11th and 12th grade students from each school type participated in the study. 108 students from Science High School, 110 students from Social Sciences High School, 108 students from Fine Arts High School and 105 students from Anatolian High School were included in the study. In general, 9th grade students consisted of 107 (24.8%), 10th grade students 106 (24.6%), 11th grade students 107 (24.8%) and 12th grade students 111 (25.8%). This shows that a balanced distribution was achieved across grade levels, with a similar number of students from each grade represented in the sample.

Data Collection Tool

Within the research framework, the Formulating Test (FT) was used to determine high school students' level of formulating skills.

Formulating Test (FT). The Formulating Test (FT), which was developed to determine the formulating levels of high school students, was created from PISA mathematical literacy problems focusing on formulating skills and made available by OECD.

Formulating is considered a fundamental component of the PISA mathematical literacy framework, and mathematical literacy problems are classified according to processes such as formulating, employing, interpreting, and reasoning (OECD, 2023). In this context, the preference for problems that directly focus on formulating skills in the creation of the FT increases the test's suitability for its purpose.

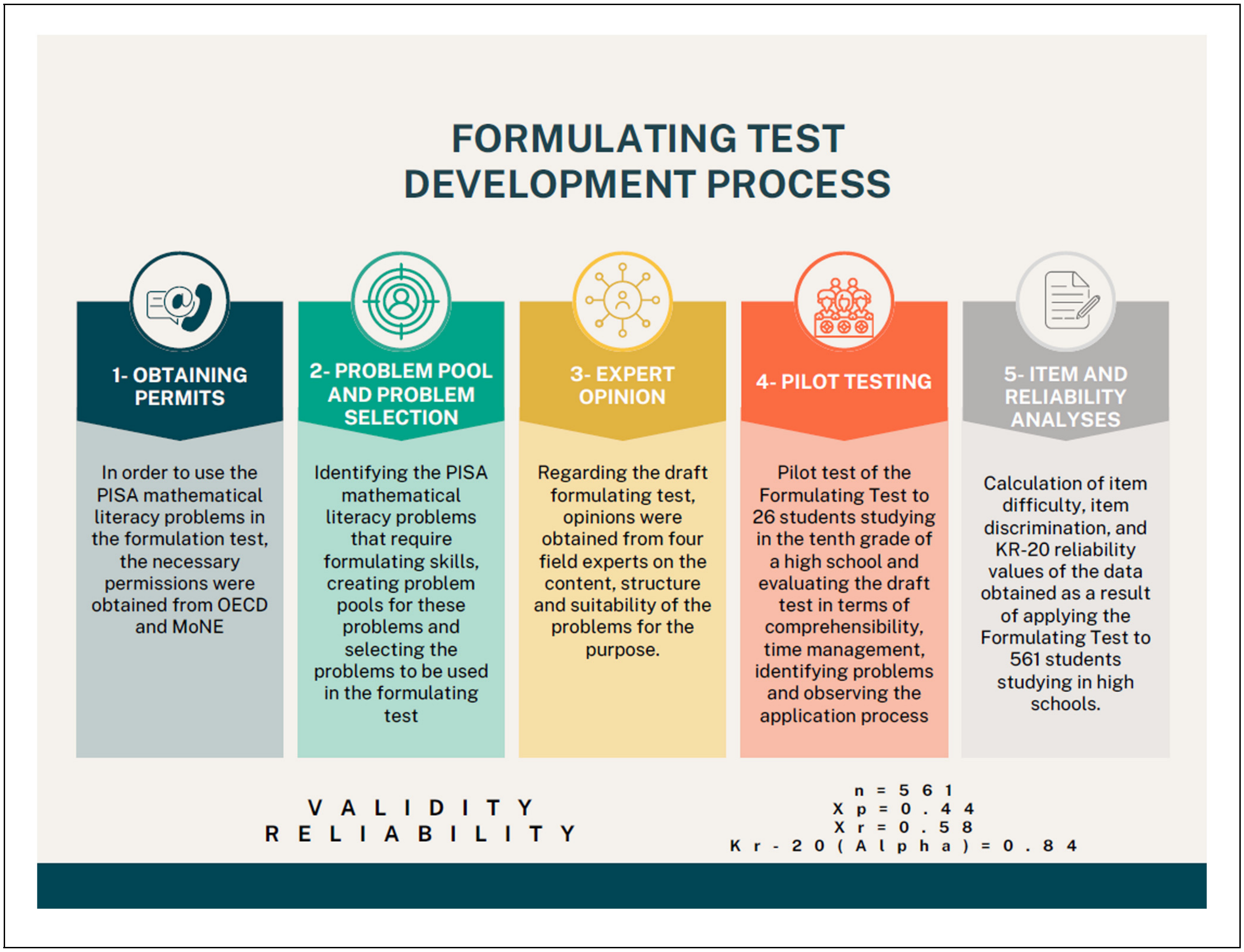

Figure 3. FT development process.

Moreover, the fact that expert teams develop PISA problems ensures that these problems meet high standards of validity and reliability. This important factor makes using these problems in the FT meaningful and valuable. The development process of the FT is visualised in Figure 3.

During the development process of the FT, the necessary permissions were first obtained from OECD for using the problems prepared by OECD in the test, and from the MoNE General Directorate of Measurement, Evaluation and Testing Services for using these problems in Turkish.

After obtaining permission, the mathematical literacy problems used by the OECD in the PISA pilot and primary applications, declassified, made publicly available, and directly related to formulating skills, were analysed in detail. As a result of these analyses, a problem pool containing appropriate problems centred on the formulating process was created.

While determining the relationship of the problems with formulating skills, the content areas and contexts in which they are located, the reports published by OECD (2013) were utilised. Information about the problems used in the draft FT—problem code, name, content area and context—is presented in Table 3.

Validity and Reliability of the Formulating Test. FT developed to determine the formulating levels of high school students consists of items scored on a binary (0–1) scale. As these items are not based on the assumption of continuous variables, they may fail to meet the normality and linearity assumptions required for traditional exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Therefore, the content validity of the items in the FT was ensured through expert opinion; construct validity was assessed using item statistics. This approach is consistent with the measurement and

Table 3. Information on the Problems in FT.

FT problem no	PISA problem code	Problem name	Content	Context
1	PM924Q02	Sauce Q2	Quantity	Personal
2	PM995Q03	Revolving Door Q3	Quantity	Scientific
3	PM995Q02	Revolving Door Q2	Space and shape	Scientific
4	PM942Q01	Climbing Mount Fuji Q1	Quantity	Societal
5	PM942Q02	Climbing Mount Fuji Q2	Change and relationships	Societal
6	PM00FQ01	Apartment Purchase Q1	Space and shape	Personal
7	PM923Q04	Sailing Ships Q4	Change and relationships	Scientific
8	PM934Q02	Ferris Wheel Q2	Space and shape	Societal
9	PM00EQ01	Faulty Players Q3	Uncertainty and data	Occupational
10	PM00EQ03	Faulty Players Q2	Uncertainty and data	Occupational
11	PM904Q04	Mp3 Players Q4	Change and relationships	Occupational
12	PM921Q02	Penguins Q2	Quantity	Scientific
13	PM921Q03	Penguins Q3	Change and relationships	Scientific
14	PM994Q01	Selling Newspapers Q1	Change and relationships	Occupational
15	PM994Q02	Selling Newspapers Q2	Change and relationships	Occupational

assessment literature on performance-based tests (DeVellis & Thorpe, 2022; Haynes et al., 1995; Lawshe, 1975).

To ensure the content validity of the draft FT, expert opinions were sought by established criteria (Cronbach & Meehl, 1955; Messick, 1995; Thorndike et al., 1991). The aim was to confirm that the test items adequately represented the measured domain. For this purpose, a form was prepared to collect expert evaluations on the suitability of the 15 problems included in the FT for measuring high school students' formulating skills. Four field experts—three professors and one assistant professor—evaluated the problems using a 5-point scale. All experts assigned the maximum score of 5, indicating that the problems were appropriate regarding purpose, content, and student level.

To determine the feature's coverage power relative to the research's scale items, an expert opinion process must be conducted (Allen & Yen, 2002; McMillan & Schumacher, 2010). This method, widely used in measurement and evaluation, is considered an important step in ensuring content validity (Haynes et al., 1995; Lawshe, 1975). In this context, the opinions of four faculty members (3 Prof. Dr., 1 Asst. Prof.) who are experts in their fields were sought to determine which dimensions the problems in the FT belong to within the framework of the formulating process model. The experts were asked to classify each problem in the FT within the framework of the formulating process model into the dimensions of Mathematical Sense-Making, Structuring and Generalising, Interpretation of Mathematical Abstractions, and Real-Life Problem Creation, and to state their reasons. Each problem was evaluated within the sub-dimension shared by at least three of the four experts. For example, Table 4 shows that in question 1, all experts classified the problem under the sub-dimension of 'Mathematical sense-making'. In contrast, in question 2, three out of four experts classified the problem under the sub-dimension of 'Interpretation of mathematical abstractions'. Table 4 presents the dimensions of the formulating process to which the problems in the FT belong, along with examples of the justifications for these classifications.

As a result of the expert evaluations conducted, it was determined that six problem fall under 'mathematical sense-making' (1, 2, 4, 12, 14, 15), four problem under 'structuring and generalising' (3, 6, 10, 13), four problem under 'interpretation of mathematical abstractions' (5, 8, 9, 11), and one problem under 'real-life problem creation' (7).

Based on expert feedback, the test was deemed suitable for its intended scope and target group, and a pilot study was subsequently conducted. The pilot was administered to 26 tenth-grade students from a high school in the Marmara region of Türkiye. Students were asked to report unclear expressions, spelling errors, or visual issues. The pilot study confirmed that the test items were comprehensible, error-free, and suitable in content and application time. As a result, the full implementation and reliability analysis stages were initiated.

Following permissions from the MoNE, the finalised FT was administered to 561 students from various high school types (Science, Anatolian, Fine Arts, and Social Sciences) across the Marmara and Aegean regions. Reliability was assessed using internal consistency analysis, which evaluates the extent to which test items measure a single construct (Johnson & Christensen, 2012). For dichotomously scored items (1 = correct, 0 = incorrect or blank), the Kuder–Richardson Formula 20 (KR-20) was used to calculate internal consistency (Kuder & Richardson, 1937; Ntumi et al., 2023).

Table 4. Classification of FT Based on Expert Opinions in the Context of Subdimensions.

FT	FT modelling sub-dimension	Example of a classification justification	Percentage of experts' consensus
1	Mathematical sense-making	The problem involves converting the given ratio into a mathematical expression. Horizontal mathematisation process	100.0
2	Mathematical sense-making	The problem involves creating a mathematical structure that accounts for the number of rotations of the revolving door and its human capacity.	75.0
3	Structuring and generalising	The problem involves converting a geometric situation in a real-life context into a mathematical structure and generalising the circle-arc relationship.	100.0
4	Mathematical sense-making	The problem involves calculating the mathematical average using data from real-life situations.	100.0
5	Interpretation of mathematical abstractions	The problem involves establishing the relationship between speed and time within an abstract mathematical structure and interpreting it within a real-life context.	75.0
6	Structuring and generalising	The problem involves discovering a practical method for calculating areas and generalising the rectangular structure.	75.0
7	Real-life problem creation	The problem involves understanding a real-life situation transformed into a problem based on an environmental context and re-expressing and formulating the problem.	75.0
8	Interpretation of mathematical abstractions	The problem involves evaluating a general, abstracted situation within the context of a rotation period and angle relationship in real-life scenarios.	100.0
9	Interpretation of mathematical abstractions	The problem involves formulating mathematical situations in real-life contexts and interpreting them in the context of mathematical generalisations.	100.0
10	Structuring and generalising	The problem involves generalising the overall situation based on two data sets and subsequently comparing the generalised structures.	100.0
11	Interpretation of mathematical abstractions	The problem involves the interpretation of generalised mathematical data.	100.0
12	Mathematical sense-making	The problem involves converting real-life situations into mathematical models using scientific information derived from assumptions.	100.0
13	Structuring and generalising	The problem involves formulating the general formula for the structure, given its mathematical equivalent, encompassing the generalisation of the structure.	100.0
14	Mathematical sense-making	The problem involves converting the pricing system into a mathematical formula to calculate the solution.	100.0
15	Mathematical sense-making	The problem involves establishing a mathematical structure for reversing the profit-sales relationship.	100.0

Note. The information in the table is based on the assessment of four experts.

Item scoring in FT followed OECD and MoNE PISA guidelines, where correct answers received 1 point, and incorrect or unanswered items received 0. The KR-20 coefficient was computed based on item difficulty, item discrimination, and overall test consistency (Başol et al., 2013). These statistics, including item difficulty and discrimination indices for each test item, along with the overall KR-20 reliability coefficient, are presented in Table 5.

When Table 5 is analysed, it is observed that the item difficulty index of the 15 problems in FT ranges from 0.10 to 0.83. According to Walsh and Betz (1995), the ideal item difficulty range is 0.10–0.90. According to Adıgüzel and Özdoğru (2013), $p \leqslant .40$ is classified as difficult, $.41 \leqslant p \leqslant .60$ as medium, and $.61 \leqslant p \leqslant .80$ as easy. Accordingly, the FT has four difficult, seven medium and four easy items. In PISA applications, the number of items with medium difficulty is high, while easy and difficult items are distributed evenly. In this respect, the problems in FT are of ideal difficulty.

Focusing on the problems in the FT within the framework of item discrimination, it is seen that the discrimination levels of the problems vary between .31 and .84. According to Towns (2014), in the item discrimination index analysis, if Discrimination Value is >0.40, the item is considered as an excellent item, and if it is between 0.20 and 0.40, it is considered as a good item. Within this framework, four items in the FT are satisfactory, and 11 are excellent in terms of discrimination.

A KR-20 reliability analysis was conducted to assess the reliability of FT. Based on the data analysis of 561 high school students, the KR-20 reliability coefficient was calculated to be 0.84 ($\alpha = .84$). According to Wallen and Fraenkel (2013), KR-20 reliability coefficient of .70 or above is required to obtain a reliable score. According to the What Works Clearinghouse (WWC) standards,

Table 5. FT Internal Consistency Analysis Results.

FT problem no	Item Difficulty Index (*Pj*)	Item Discrimination Index (*Pj*)
1	0.83	0.36
2	0.58	0.56
3	0.10	0.31
4	0.73	0.49
5	0.42	0.72
6	0.61	0.71
7	0.12	0.35
8	0.62	0.53
9	0.24	0.54
10	0.42	0.80
11	0.16	0.37
12	0.46	0.84
13	0.42	0.52
14	0.47	0.78
15	0.48	0.80

Note. Number of items = 15; number of participants (n) = 561; KR-20 (alpha value) (α) = .84; average difficulty value (Xp) = 0.44; average discrimination value (Xr) = 0.58.

the internal consistency coefficient for educational studies should be greater than .60 (WWC, 2022). Atalay et al. (2015) stated that a KR-20 value between 0.60 and 0.70 is considered acceptable, values between 0.70 and 0.90 are considered good, and values above 0.90 are considered excellent. In this framework, the reliability value obtained is nearly perfect. As a result of the analyses, it is evident that FT is a valid and reliable test for measuring students' formulating levels.

Data Analyses

The study used FT to determine high school students' level of formulating skills. The students' responses to the FT were scored as 1 for correct answers and 0 for incorrect or blank answers, following the scoring method used for PISA problems. The data obtained in this context were transferred to the IBM Statistical Package for Social Sciences for Personal Computers (SPSS 26.0) programme, and statistical analyses were carried out within the research problem framework.

Item response theory and Bloom's cognitive domain theory were considered in evaluating the FT scores. A criterion-based evaluation approach was adopted to determine the formulating levels, and achievement levels were classified according to objective criteria (Anderson & Krathwohl, 2001; Lok et al., 2016). In this context, high-level students (11–15 points) are expected to show analysis, synthesis and evaluation skills by solving complex problems. Medium-level students (6–10 points) generally answer easy and medium problems correctly and show deficiency in complex problems; this group shows competence in comprehension and application steps. Low-level students (0–5 points) are inadequate in formulating skills and show deficiencies in fundamental knowledge and understanding.

Before the statistical analyses, the data were examined for compliance with normal distribution. This examination aims to determine whether the tests' assumptions are met and to obtain more valid results by using appropriate alternative tests in the event of possible violations (Can, 2014). Violating the normality assumption may undermine the validity of the results by reducing statistical power (Wells & Hintze, 2007). In educational research, this assumption is assessed using descriptive statistics, graphical methods, and hypothesis tests (Demir et al., 2016). The study examined the normality assumptions by evaluating skewness and kurtosis coefficients using descriptive methods. According to Tabachnick and Fidell (2013), coefficients between −1.5 and + 1.5 indicate that the data are normally distributed. In the study, before determining whether there was a significant difference in formulating skills across school type, grade level, and gender variables, the skewness and kurtosis values of the data groups were examined. It was observed that all groups fell within this range, and the assumption of normality was accepted.

The analysis was conducted in accordance with the research's sub-problems. First, a repeated-measures ANOVA was conducted, and Mauchly's Sphericity test was performed to examine performance differences across the four sub-dimensions of the formulating model, assuming normality. A Bonferroni correction was used to adjust for multiple comparisons. ANOVA was preferred to determine differences across school type and grade level variables, as the data were normally distributed; however, because the variances were non-homogeneous, Welch ANOVA was used, followed by the Games-Howell post hoc test. Finally, an independent-samples t-test was used on data that met the normality assumption to determine whether formulating skills differed by gender. This structure ensured the analysis was conducted systematically and consistently. Besides, a three-way between-groups factorial ANOVA was conducted to examine the main and interaction effects of school type, grade level, and gender on students' formulating skills.

In addition, the focus was on the effect size in determining the extent to which the type of school and grade level significantly affect the formulating skills of high school students. Effect size is a statistical value that expresses the strength of the relationship between variables (Hedges, 2008) and shows the extent to which the results obtained from the sample deviate from the null hypothesis (J. Cohen, 1994). Today, many educational journals require reporting of effect size in addition to

Table 6. Multiple ANOVA Analysis Results According to Sub-Dimensions in the Formulating Model.

Sub-dimensions	*N*	*M*	*SD*	Source	Type III sum of squares	*df*	Mean square	*F*	*p*	η^2 *p*
Mathematical sense-making	431	3.50	1.87	Greenhouse-Geisser	2,634.22	2.01	1,310.12	997.29	.00*	.70
Structuring and generalising	431	1.20	1.04							
Interpretation of mathematical abstractions	431	1.36	1.36							
Real-life problem creation	431	0.08	0.27							

*$p < .05$.

t- or *F*-statistics (Huberty, 2002). Therefore, the effect size was calculated in this study. In one-way ANOVA analyses, the eta squared (η^2) value is generally used to assess effect size (Lakens, 2013). Accordingly, the eta squared value was calculated in this study. J. Cohen (1988) determined the effect size limits for the *f* statistic as small (0.10), medium (0.25) and large (0.40). When these values are converted into eta squared, they correspond to 0.01 (small), 0.06 (medium) and 0.14 (large) effect sizes, respectively (Green & Salkind, 2005).

Ethical Considerations

This research was conducted with the approval from Ministry of National Education and the relevant school administration. The study was conducted with the permission of the Ethics Committee of the Faculty of Science and Engineering at Balikesir University (Protocol No: E-19928322-302.08.01-32076).

Before implementing the applications, participants' parents were informed and provided consent via the consent form. The students' statements regarding their voluntary participation in the research were obtained, and they were informed that they could withdraw from the research at any time without giving any reason.

To minimise potential risks, it has been stated that the research was not conducted for evaluation or grading purposes and will not affect students' academic standing. Data was collected anonymously to ensure participant confidentiality. This study aims to contribute to a better understanding of students' formulation skills in mathematics education. In this regard, it is believed that the expected scientific and social benefits of the research outweigh the risks that participants may encounter.

Findings

The findings section is structured according to the four sub-problems examined in the research. The results are presented under subheadings, with the relevant analysis methods applied for each sub-problem. This aims to ensure that both the statistical findings and their interpretations are consistent with the problem statement.

Finding on the Comparison of High School Students' Success Achievement Levels According to the Formulating Process Framework Sub-Dimensions?

Firstly, within the scope of RQ1, it was examined whether there was a significant difference in the success of high school students in the sub-dimensions of the formulating model. The scores obtained by students in FT for each sub-dimension were calculated. Then, multiple ANOVA analyses were used to determine whether there were significant differences in the scores of high school students across the sub-dimensions. The results of the multiple ANOVA analyses are presented in Table 6, along with descriptive findings:

A repeated-measures ANOVA was conducted to examine significant differences in the scores of high school students across the sub-dimensions of the formulating model. Considering Mauchly's sphericity test, a significant difference was found ($\chi^2(5) = 297.77$, $p < .001$). Within this framework, the results were reported with the Greenhouse-Geisser correction. The analysis results indicate a significant difference among the sub-dimensions ($F(2.01, 864.59) = 997.29$, $p < .001$, η^2 $p = .70$). In this context, the sub-dimensions that differed significantly were identified using Bonferroni-corrected pairwise comparisons following the repeated-measures ANOVA. The results are presented in Table 7.

When examining the descriptive statistics conducted to determine significant differences in high school students' scores across sub-dimensions, it was observed that students scored highest on the mathematical sense-making sub-dimension ($M = 3.50$, $SD = 1.87$) and lowest on the real-life problem creation sub-dimension ($M = 0.08$, $SD = 0.27$). When Bonferroni-corrected pairwise comparisons were examined, it was concluded that sense-making scores were significantly higher than all other sub-dimensions ($p < .001$). Furthermore, the interpretation of mathematical abstractions sub-dimension

Table 7. Sub-Dimension Pairwise Comparisons (Bonferroni Corrected).

Factor (*I*)–Factor (*J*)	Mean difference (*I-J*)	*SE*	*p**	95% CI lower bound	95% CI upper bound
1–2	2.30	0.07	.000	2.12	2.48
1–3	2.14	0.07	.000	1.94	2.33
1–4	3.42	0.09	.000	3.19	3.65
2–3	−0.17	0.05	.002	−0.29	−0.04
2–4	1.12	0.05	.000	0.99	1.24
3–4	1.28	0.05	.000	1.15	1.42

Note. The Bonferroni correction has been applied. 1 = mathematical sense-making, 2 = structuring and generalising, 3 = interpretation of mathematical abstractions, 4 = real-life problem creation, SE = standard error, CI = confidence interval.
*$p < .05$.

Table 8. ANOVA Analysis Table of FT Scores According to School Type.

School	*N*	*M*	*SD*	Source of variance	Sum of squares	*df*	Mean square	*F*	*p*	Welch *p*	Sig. difference
Science High School (1)	108	8.69	3.24	Between Gr.	2,582.66	2	860.56	120.93	.00*	.00*	1–2
Social Sciences High School (2)	110	5.82	2.80	Whitin Gr.	3,038.54	427	7.12				1–3
Fine Arts High School (3)	108	2.31	1.68	Total	5,620.20	430					2–3
Anatolian High School (4)	105	7.78	2.70								4–2
											4–3

*$p < .05$.

was found to be significantly higher than structuring and generalising (p = .002). These findings indicate that students are particularly strong in the mathematical sense-making dimension but perform relatively poorly in the real-life problem creation dimension.

Findings on the Comparison of High School Students' Formulating Skills According to School Types

Within the scope of RQ2, it was examined whether high school students' formulating skills differed by school type. Data from the FT administered to students at Anatolian High School, Fine Arts High School, Science High School, and Social Sciences High School were analysed to determine whether the levels differed significantly by school type. In this regard, ANOVA findings comparing students' FT scores by school type are presented in Table 8.

When the ANOVA results in Table 8 were analysed, it was found that the variances of the students' FT scores across school types were not homogeneously distributed (p = .00, $p < .05$). Therefore, the Welch test results were considered. The significance value [$.F(2, 427) = 120.93$, p = .00, $p < .05$] obtained in the ANOVA revealed that there was a significant difference in formulating skills according to school type. Games–Howell post hoc test was applied to determine which groups this difference was between.

According to the analyses, significant differences were found between Science High School and Social Sciences and Fine Arts High Schools; between Social Sciences High School and Fine Arts and Anatolian High Schools; and between Anatolian High School and Fine Arts High School. When the mean FT scores were analysed, it was observed that Science High School Students (M = 8.69) scored higher than Social Sciences (M = 5.82) and Fine Arts High School students (M = 2.31). Similarly, it was determined that Anatolian High School Students (M = 7.78) obtained higher scores than both school types, and that the scores of Social Sciences High School students were higher than those of Fine Arts High School students. In this context, the formulating skills of Science and Anatolian High School students were higher than those of Social Sciences and Fine Arts High School students.

When the mean scores were evaluated by school type, the formulating skill levels of students from Science High School, Anatolian High School, and Social Sciences High School were classified as 'medium'. In contrast, the levels of Fine Arts High School students were classified as 'low'. These findings show that the development of formulating skills can vary significantly across school types and that school type is an important factor in developing these skills. In addition, the calculated eta squared value (η^2 = .46), which is above .14, indicates a high effect size level. This supports the idea that school type is a determinant variable in formulating skills.

Table 9. ANOVA Analysis Table of FT Scores According to Grade.

Grade	*N*	*M*	*SD*	Source of variance	Sum of squares	*df*	Mean square	*F*	*p*	Welch *p*	Sig. difference
9th	107	4.60	3.01	Between Gr.	4,554.84	2	151.61	12.53	.00*	.00*	10–9
10th	106	6.45	3.45	Whitin Gr.	5,165.35	427	12.10				11–9
11th	107	6.00	3.75	Total	5,620.20	430					12–9
12th	111	7.44	3.64								12–11

*$p < .05$.

Findings on the Comparison of High School Students' Formulating Skills According to Grade

Within the scope of RQ3, it was examined whether high school students' formulating skills differed by grade level. The results of the ANOVA analysis, which aim to determine significant differences in formulating skills among high school students by grade level, are presented in Table 9.

According to the ANOVA results in Table 9, the variances of high school students' FT scores by grade level were not homogeneously distributed ($p = .00$, $p < .05$). Therefore, the Welch test was applied. The results of the analysis showed that the significance level was below .05 [$F(2, 427) = 10.71$, $p = .00$, $p < .05$]. These results reveal a significant difference between the grade levels regarding students' formulating skills.

As a result of the Games–Howell post hoc analysis, which determined which grade levels the difference was between, significant differences were found between the 9th and 10th, and 11th and 12th grades, in favour of the upper grades, respectively. In addition, a significant difference was found between the 11th and 12th grades in favour of the 12th. Generally, the highest level of formulating skill was observed in the 12th grade, and the lowest level in the 9th grade. This finding shows that students' formulating skills improve as the grade level increases.

When the mean FT scores according to grade level were analysed, it was found that 10th, 11th and 12th grade students ($M > 5$) had 'medium' formulating skills, while 9th grade students ($M < 5$) had 'low' formulating skills. In addition, the eta squared ($\eta^2 = .08$) value calculated for the grade level variable was between .06 and .14, indicating a medium effect size. These findings indicate that grade level may be a factor to consider in developing and formulating skills.

Findings on the Comparison of High School Students' Formulating Skills According to Gender

Within the scope of RQ 4, it was analysed whether students' formulating skills differed by gender. To determine whether there is a significant difference by gender, the independent-samples *t*-test, a parametric analysis method, was applied. Table 10 presents the independent-sample *t*-test analysis results:

Table 10. *T* Test Table of FT Scores According to Gender.

Group	*N*	*M*	*SD*	*df*	*t*	*p*
Female	226	6.29	3.60	429	.92	.36
Male	205	5.97	3.64			

$p > .05$.

As a result of the *t*-test analysis conducted to determine whether there is a significant difference in the FT score averages of high school students by gender, the *p*-value was greater than .05 ($t(429) = 0.92$, $p = .13$, $p > .05$). This result shows that the gender variable does not have a statistically significant effect on formulating skills. In addition, when the participants' FT scores were examined by gender, it was observed that the average score for female students ($M = 6.29$) and for male students ($M = 5.97$) were at a moderate level and quite close to each other. This finding indicates that gender is not a determining factor in formulating skills.

Findings on the İnteraction Effect of School Type, Grade Level, and Gender Variables on Formulating Skills

Within the scope of RQ 5, the interaction effect of school type, grade level, and gender on formulating skills was examined. To reveal the relationships among variables, reduce the probability of Type 1 error, and determine whether the combined effects of school type, grade level, and gender on secondary school students' formulation skills are significant, a three-way between-groups factorial ANOVA was utilised. The results of the between-groups factorial ANOVA are presented in Table 11:

According to the factorial ANOVA results, school type was found to have a significant effect on students' formulation skills, $F(3, 399) = 148.14$, $p < .001$, $\eta^2 = .53$. Bonferroni-corrected pairwise comparisons revealed that all differences between school types were statistically significant ($p < .05$). Additionally, the type of school

Table 11. Factorial ANOVA Results for FT Scores by School Type, Grade Level, and Gender.

Source	Type III sum of squares	*df*	Mean square	*F*	*P*	Partial eta squared
Corrected model	3,379.48[a]	31	109.02	19.41	.00*	.60
Intercept	15,852.94	1	15,852.94	2,822.90	.00*	.88
School type	2,495.75	3	831.92	148.14	.00*	.53
Grade	422.28	3	140.76	25.07	.00*	.16
Gender	2.46	1	2.46	.44	.51	.00
School type × Grade	254.26	9	28.25	5.03	.00*	.10
School type × Gender	25.02	3	8.34	1.49	.22	.01
Grade × Gender	21.28	3	7.09	1.26	.20	.01
School type × Grade × Gender	70.32	9	7.81	1.39	.19	.03
Eror	2,240.72	399	5.62			
Total	21,840	431				
Corrected total	5,620.20	430				

Note. [a]$R^2 = .601$ (adjusted $R^2 = .570$).
*$p < .001$.

has been found to have a significant effect size, indicating that the kind of school explains a substantial portion of the variance in students' formulating skills. Similarly, class level was found to have a significant effect on formulation skills, $F(3, 399) = 25.07$, $p < .001$, $\eta^2 = .16$. In contrast, it was concluded that gender had no significant effect on formulation skills, $F(1, 399) = 0.44$, $p = .51$, $\eta^2 = .00$.

The analysis results indicate that the interaction between school type and grade level is statistically significant, $F(9, 399) = 5.03$, $p < .001$, $\eta^2 = .10$. This finding reveals that secondary school students' formulation skills differ across school types by grade level. In summary, while secondary school students' formulation skills differ by school type and grade level, they do not vary by gender.

Discussion

Formulating skill is employed in the processes of mathematical expression of situations in a real-life context (OECD, 2023), abstracting these expressions by transforming them into a systematic structure (Tall, 2013), interpreting and applying abstract concepts to real-life context (Blum & Niss, 1991) and transforming these applications into problem situations. Formulating skill is an important skill not only in mathematical literacy problems but also in the mathematisation process in realistic mathematics education (Artut & Bal, 2016; Demirdöğen & Kaçar, 2010; Zulkardi, 2000), mathematical modelling process (Ang, 2010; Berry & Houston, 1995; Blomhøj & Jensen, 2007; Mason, 1988; Stillman et al., 2007; Tuna et al., 2013; Voskoglou, 2006) and problem-solving process (Temel & Altun, 2022). So, what is the status of high school students in formulating, which is critical for mathematical skills and concepts?

The research revealed that high school students' formulating skills have a multidimensional structure and that students demonstrate different levels of performance in the sub-dimensions of this skill. Students' higher scores in the mathematical sense-making dimension indicate that they are relatively more successful in understanding mathematical situations and establishing connections between mathematics and real life. However, the lowest performance is observed in the real-life problem-creation dimension, indicating that students have difficulty transforming their mathematical knowledge into real-life contexts and creating original problems. This supports the finding that formulating was the process in which students struggled most in their mathematical literacy in PISA 2012 and 2022 (OECD, 2014, 2023). Furthermore, another notable finding is that students were more successful at interpreting mathematical abstractions than at structuring and generalising. This parallels the findings of the study by Ellis et al. (2022), which found that students struggled when asked to create a new generalisation; however, they were more successful at relating an existing generalisation to different contexts or extending it to a new domain. Similarly, these results are consistent with Mason's (2005) study, which revealed that students experience difficulties in producing generalisations.

The findings from the FT indicate that students' formulating skills in the 10th, 11th, and 12th grades at Science High Schools, Anatolian High Schools, and Social Sciences High Schools are at a medium level. In contrast, those of 9th-grade students and students at Fine Arts High Schools remain low. These results once again indicate that students generally experience difficulties in the formulating process. Similar findings from studies conducted in Türkiye (Karaduman et al., 2023; Özgen, 2019; Ülger et al., 2020) support the difficulty in the formulating process. At the same time, the low correct-answer rates on some problems in the FT (problems 3, 7, 11, etc.) also indicate that students struggle, especially in situations requiring higher-level mathematisation. In this context, it is necessary to develop and expand teaching practices that focus on the formulating process to improve students' mathematical literacy.

Türkiye's PISA 2022 results show significant differences in mathematics literacy scores across high school types. Science High Schools achieved the highest performance, while Social Sciences High Schools, Anatolian High Schools, Fine Arts High Schools, and Vocational and Technical Anatolian High Schools showed lower performance, respectively (MoNE, 2023). This situation is consistent with the findings reported by Suna et al. (2020) on differences between school types using PISA 2003–2018 data. The results of the study show that Science High School students are more successful in formulating skills than students from other school types. In contrast, Fine Arts High School students are at the lowest level, consistent with PISA results and reports by Suna et al. (2020). This consistency suggests that school type is a structural determinant of mathematical performance, particularly in formulating skills.

In addition, the fact that formulating skills remain at a medium level even in schools with longer mathematics lesson times and greater programme intensity indicates that this skill is not sufficiently reinforced in teaching. In other words, attributing the difference between school types solely to curriculum intensity is insufficient. The fact that even students in Türkiye's leading secondary schools have only medium-formulating skills supports the criticism, long emphasised in international assessments such as PISA (OECD, 2014, 2023), that mathematical modelling and mathematisation processes are not sufficiently addressed in teaching. Furthermore, despite the increase in content intensity, the continued limitation of students' ability to transform real-life situations into mathematical structures indicates that teaching practices remain content-focused rather than process-focused.

The low performance of Fine Arts High School students is consistent with the two fundamental problems highlighted by Döğer and Gayretli (2019) regarding this type of school: (1) the limited time allocated to mathematics teaching in this type of school, and (2) students' negative attitudes towards numerical subjects. However, the findings go beyond this situation and, when considered alongside studies (Yu & Li, 2022) showing that representation and spatial reasoning skills are critical for both artistic production and formulating processes, suggest that Fine Arts High School students may have a potential cognitive advantage. The fact that this advantage does not translate into performance indicates that the teaching design specific to this type of school does not activate this potential. Therefore, it is thought that the low performance of Fine Arts High School students is related not only to inadequate mathematics education but also to the structure of current teaching models, which do not incorporate interdisciplinary cognitive approaches. This criticism coincides with the reasons why STEAM-based approaches are gaining increasing importance in the international literature (Preciado-Babb et al., 2013).

The study revealed that formulating skills increase with grade level, with 9th-grade students demonstrating the lowest performance and 12th-grade students the highest. This finding is consistent with Wildani's (2020) findings, which indicate that 9th-grade students in particular experience difficulty in converting verbal problems into mathematical structures. This increase in formulating across grade levels can be attributed to the development of students' cognitive capacity and academic competence. When evaluated in the context of Piaget's cognitive development theory, 9th-grade students are transitioning from concrete to formal operations. Therefore, it is expected that students will struggle with higher-level skills such as abstraction, representation creation, and reaching generalisations, such as formulating (Piaget, 1952). However, this explanation is open to criticism because it assumes that cognitive development progresses with age. This is because skills such as mathematical modelling, reasoning, and problem solving are shaped not only by age but also by educational experience and teaching quality (Harel & Sowder, 2007; Schoenfeld, 1992). Therefore, the increase at the grade level should be associated not only with developmental maturation but also with students gaining more experience in mathematical modelling, reasoning, representation, and problem solving in later years. In this context, the low performance at the 9th-grade level can be seen as a reflection of both the cognitive transition period and the lack of adequate learning opportunities to develop mathematical modelling, problem-solving, representation, and formulating skills at this grade level.

The finding that gender has no significant effect on formulating skills indicates that this skill is independent of gender. This finding is consistent with Hyde's (2005) 'gender similarity hypothesis', which suggests that gender differences in mathematical ability and similar cognitive domains are generally low. Furthermore, the study by Else-Quest et al. (2010), which provides empirical evidence that gender has no significant effect on mathematical performance in many contexts, is consistent with the present findings. In this respect, the study aligns with the literature, which shows that formulating skills are more closely related to learning experiences and teaching processes than to individual differences.

Conclusion

The results of this study indicate that high school students' formulating skills are generally inadequate, particularly in the processes of structuring, generalising and real-life problem creation. It was found that formulating skills are independent of gender but vary with factors such as grade level and school type. It was observed that school types with programmes that place greater emphasis on mathematics provide more support for formulating skills. In contrast, in school types where mathematics teaching is limited, these skills remain at a low level. Results at the grade level show that students' formulating skills develop over the years, but 9th-grade students require more support in developing them.

Recommendations

Recommendations for Educational Applications

The research findings indicate that current teaching programmes are not sufficiently effective in developing formulating skills. Therefore, teaching programmes need to be restructured to include more applications and activities that enable students to transform real-life situations into mathematical structures and generalise them. It is critically important to systematically integrate activities involving representation creation, modelling, problem setting, and structured discussion processes into the programmes. Such activities will not only serve as cognitive tasks that develop students' formulating skills but will also help them gain experience in interdisciplinary problem-solving.

Interventions aimed at developing formulating skills within the framework of findings should not be limited to adjustments to the teaching programme. Incorporating mathematical modelling, problem-solving, realistic mathematics education, and STEM applications into the curriculum is not sufficient on its own. Teaching processes must be redesigned to activate students' representation and transformation skills, enabling them to convert everyday situations into mathematical structures. STEM activities and mathematical modelling are considered critical, particularly for contributing to the abstraction and structuring processes students need when dealing with complex problems.

Schools with limited mathematics teaching hours, such as Fine Arts High Schools, should incorporate interdisciplinary teaching practices that integrate art and mathematics. Developing STEAM-focused teaching designs that consider interdisciplinary cognitive levels for such schools can enhance students' formulating skills. Such activities can advance students' formulating skills, particularly their representational and spatial reasoning abilities. Furthermore, developing pedagogical strategies to transform negative attitudes towards these subjects in schools with few numerical courses can improve students' performance in mathematical tasks.

The study's results indicate that 9th-grade students require more support in their formulating skills than students in other grades. In this context, students starting high school must be supported in their mathematics lessons with teaching strategies that strengthen the processes of establishing mathematical relationships, abstraction and mathematisation. Supporting students through modelling activities that facilitate their transition from concrete situations to abstract concepts and incorporating applications that increase representation diversity into teaching programmes can support the development of the formulating process.

Recommendations for Policy and Programme Development

Differences in the mathematical content between school types at the secondary level can be one factor affecting students' formulating skills. Particularly in schools with intensive mathematical content, students' better performance demonstrates the impact of the quality of the teaching programme. Therefore, policymakers need to develop an inclusive programme that focuses on students' mathematical literacy and formulating skills in these schools, considering the needs of different school types. When redesigning the teaching process for schools with less mathematical content, programme developers are advised to focus on strengthening interdisciplinary integration (mathematics-arts, vocational subjects-mathematics) within a mathematical context. In this regard, it is recommended that teaching content be supported by STEAM-based applications and interdisciplinary connections to develop formulating skills.

The findings also convey a message that teacher training needs to be reviewed. One of the fundamental duties of teachers is to guide students and involve them in the learning process to develop their learning skills (Fullan et al., 2017). In this regard, enhancing teachers' competence in pedagogical methods that support students' formulating skills, such as mathematical representation, modelling, mathematisation, problem formulation, and higher-order thinking, may enable more effective development of students' formulating skills in classroom applications. Therefore, it is recommended that in-service training programmes for teachers be restructured to include comprehensive educational content focused on exemplary practices and interdisciplinary connections that will develop students' formulating skills.

Recommendations for Researchers

One of the key findings of this study is that students experience difficulties in their formulating skills. In this context, it is recommended that future research focus on the sources of the difficulties students encounter in formulating problems. Within this scope, the reasons for the difficulties students experience in the formulating stage when transferring complex problems into the mathematical world and converting generalisations and abstractions into real-life applications can be investigated in depth through qualitative studies.

Another issue that needs to be addressed in the future is how to develop students' formulating skills. In this context, researchers can examine the impact of processes that can be incorporated into teaching, such as mathematical literacy problems, modelling-based activities, and problem-solving strategies, on the development of formulating skills. In this regard, it is recommended that modelling activities based on real-life scenarios and instructional processes developed within the framework of different problem-solving strategies be systematically implemented in the early years of high school and that longitudinal studies examining the effects of different teaching approaches on formulating be conducted. Such studies will reveal the development process of the skill and more clearly identify effective intervention strategies.

Another issue that has emerged in the research and awaits further examination in future studies is the structural and pedagogical mechanisms underlying differences between school types. There is a need not only for reporting performance-based differences but also for explanatory models and model development studies that relate these differences to their causes.

Limitations

The FT developed within the framework of the study was constructed with questions that measure formulating skills within the scope of PISA mathematical literacy problems. However, the fact that the vast majority of the problems in the FT cover processes 'Mathematical sense-making', 'Structuring and generalising' and 'Interpretation of mathematical abstractions' of the formulating model in Figure 2 is considered a methodological limitation of the study. The fact that there is only one problem in the real-life problem creation dimension is another limitation of the study. For these reasons, it is suggested that tests that will address the formulating process more broadly and include different problem types should be developed in future research.

Another limitation of the study is that the participants were only students from Science High School, Social Sciences High School, Fine Arts High School and Anatolian High School. This may limit the generalizability of the findings to students studying in different types of high schools. Future studies with a broader range of students may provide a more comprehensive assessment of formulation skills.

The study examined formulating skills measured at a specific time, and this cross-sectional approach provides a limited perspective on students' developmental processes over time. Therefore, conducting longitudinal studies in future studies will help analyse the changes in students' formulating skills more comprehensively. Studies based on long-term follow-up may also provide more in-depth information about which learning processes develop these skills.

Finally, the data collection tools and methods used in the study also have some limitations. The data were collected only through FT. In future studies, using more diverse data collection methods could enable formulating skills to be assessed through written and oral performances. In addition, studies that combine quantitative and qualitative research methods could provide a more in-depth understanding of students' formulating skills.

Acknowledgements

This study utilized generative tools, such as Grammarly, to enhance language proficiency.

ORCID iD

Hasan Temel https://orcid.org/0000-0003-4532-0529

Ethical Considerations

Prior to the research, the necessary application permits were obtained in accordance with the letter numbered E-49614598-605.01-25591877 from the Ministry of National Education's Strategy Development Directorate. Within the scope of this permit, approval was obtained by applying to the relevant school administration. The study was also approved by the Ethics Committee of the Faculty of Science and Engineering at Balıkesir University (Ethics Committee Protocol No: E-19928322-302.08.01-32076).

Consent to Participate

Informed consent was obtained from the parents of the students participating in the study; the students themselves also declared that they participated voluntarily. Participants were informed that they could withdraw from the study at any time without providing any reason. During the study, participants' identity information was kept confidential in accordance with ethical guidelines, and the data obtained were used solely for scientific purposes.

Consent for Publication

The author takes full responsibility for the claims presented in this article and has given explicit permission for its submission to Sage Open.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Research Fund of Balıkesir University. Project Number: 2021/003.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

The data that support the findings of this study are available from the author upon reasonable request.

References

Adıgüzel, O. C., & Özdoğru, F. (2013). Development of an academic achievement test for common compulsory foreign language I course of universities. *Trakya University Journal of Education*, *3*(2), 1–11.

Aguirre, J. M., Turner, E. E., McVicar, E., McDuffie, A. R., Foote, M. Q., & Carll, E. (2024). Mathematizing the world: A routine to advance mathematizing in the elementary classroom. *The Journal of Mathematical Behavior*, *76*, 101196. https://doi.org/10.1016/j.jmathb.2024.101196

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory*. Waveland Press.

Altun, M. (2015). *Eğitim Fakülteleri ve Matematik Öğretmenleri için Liselerde matematik öğretimi* [*Teaching mathematics in high schools for faculties of education and mathematics teachers*] (8th ed.). Aktüel Alfa Akademi Basım Yayım Dağıtım.

Altun, M. (2020). *Matematik okuryazarlığı el kitabı* [*Mathematical literacy handbook*]. Aktüel Alfa Akademi Yayıncılık.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: Complete edition*. Addison Wesley Longman.

Ang, K. C. (2010, December 17–21). *Teaching and learning mathematical modelling with technology* [Conference session]. 15th Asian Technology Conference in Mathematics (2010), Kuala Lumpur, Malaysia. https://repository.nie.edu.sg/server/api/core/bitstreams/3e930bcf-707e-480f-acd7-9d17c2037cd6/content

Artut, P. D., & Bal, A. P. (2016). An application example on the realistic mathematics education. *International Journal of Social Sciences and Education Research*, *2*(4), 1248–1255. https://doi.org/10.24289/ijsser.279000

Atalay, N. Ş., Akkaya, N., & Şahin, F. (2015). The psychometric properties of the Turkish version of revised 2011-osteoporosis knowledge test. *Turk Osteoporoz Dergisi* [*Turkish Journal of Osteoporosis*], *21*(3), 127–131.

Başol, G., Çakan, M., Kan, A., Özbek, Ö. Y., Özdemir, D., & Yaşar, M. (2013). *Eğitimde ölçme ve değerlendirme* [*Measurement and evaluation in education*]. Pegem Akademi Yayınları

Berry, J., & Houston, K. (1995). *Mathematical Modelling*. J. W. Arrowsmith Ltd.

Blomhøj, M., & Jensen, T. H. (2007). What's all the fuss about competencies? Experiences with using a competence perspective on mathematics education to develop the teaching of mathematical modelling. In W. Blum, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education: The 14th ICMI study* (pp. 45–56). Springer US.

Blum, W., & Leiss, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modelling: Education, engineering and economics—ICTMA 12* (pp. 222–231). Horwood.

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects—State, trends and issues in mathematics instruction. *Educational Studies in Mathematics*, *22*(1), 37–68. https://doi.org/10.1007/BF00302716

Cai, J., & Howson, G. (2012). Toward an international mathematics curriculum. In M. A. Clements, A. Bishop, C. Keitel, J. Kilpatrick, & F. Leung (Eds.), *Third international handbook of mathematics education* (pp. 949–974). Springer. https://doi.org/10.1007/978-1-4614-4684-2_29

Can, A. (2014). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi* [*Quantitative data analysis in scientific research process with SPSS*]. Pegem Akademi.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997–1003.

Cohen, L., Manion, L., & Morrison, K. (2005). *Research methods in education* (5th ed.). Routledge Falmer.

Coleman, J. S. (1990). *Foundations of social theory*. Harvard University Press.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Demir, E., Saatçioğlu, Ö., & İmrol, F. (2016). Examination of educational researches published in international journals in terms of normality assumptions. *Current Research in Education*, *2*(3), 130–148.

Demirdöğen, N., & Kaçar, A. (2010). The effect of realistic mathematics education approach on the student's success of teaching fraction concept in 6th class Erzincan University. *Journal of Education Faculty*, *12*(1), 56–74.

DeVellis, R. F., & Thorpe, C. T. (2022). *Scale development: Theory and applications*. Sage.

Dewantara, A. H., Zulkardi, Z., & Darmawijoyo, D. (2015). Assessing seventh graders' mathematical literacy in solvıng PISA-like tasks. *Journal on Mathematics Education*, *6*(2), 39–49.

Döğer, D., & Gayretli, Ş. (2019). Investigation of fine arts high school music department students' achievement in quantitative courses and their opinions on these courses. *Eğitim ve Toplum Araştırmaları Dergisi*, *6*(2), 336–354.

Doorman, L. M., & Gravemeijer, K. P. E. (2009). Emergent modeling: Discrete graphs to support the understanding of change and velocity. *ZDM*, *41*, 199–211. https://doi.org/10.1007/s11858-008-0130-z

Edo, S. I., Hartono, Y., & Putri, I. I. R. (2013). Investigating secondary school students' difficulties in modeling problems PISA-Model level 5 and 6. *Journal on Mathematics Education*, *4*(1), 41–58.

Ellis, A. B., Lockwood, E., Tillema, E., & Moore, K. (2022). Generalization across multiple mathematical domains: Relating, forming, and extending. *Cognition and Instruction*, *40*(3), 351–384. https://doi.org/10.1080/07370008.2021.2000989

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127. https://doi.org/10.1037/a0018053

Ferri, R. B. (2006). Theoretical and empirical differentiations of phases in the modelling process. *ZDM*, *38*, 86–95. https://doi.org/10.1007/BF02655883

Fowler, F. J. (2014). *Survey research methods*. Sage.

Fullan, M., Quinn, J., & McEachen, J. (2017). *Deep learning: Engage the world change the world*. Corwin Press.

Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh—Analyzing and understanding the data*. Pearson.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. *Second Handbook of Research on Mathematics Teaching and Learning*, *2*, 805–842.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*(3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238

Hedges, L. V. (2008). What are effect sizes and why do we need them? *Child Development Perspectives*, *2*(3), 167–171. https://doi.org/10.1111/j.1750-8606.2008.00060.x

Hedström, P. (2005). *Dissecting the social: On the principles of analytical sociology*. Cambridge University Press.

Huberty, C. J. (2002). A history of effect size indices. *Educational and Psychological Measurement*, *62*(2), 227–240. https://doi.org/10.1177/0013164402062002002

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Johnson, B., & Christensen, L. (2012). *Educational research quantitative, qualitative, and mixed approaches* (4th ed.). Sage.

Karaduman, B., Arslan, Ç., Broutın, M. S. T., & Ezentaş, R. (2023). Investigation of the mathematical literacy problems of special talented students according to the art, music and numerical talents. *Turkish Journal of Educational Studies*, *10*(2), 193–220. https://doi.org/10.33907/turkjes.1150966

Kennedy, T. J., & Odell, M. R. (2014). Engaging students in STEM education. *Science education international*, *25*(3), 246–258. https://files.eric.ed.gov/fulltext/EJ1044508.pdf

Köroğlu, H., & Yeşildere, S. (2004). Learner achievement effect of the multiple intelligences theory based teaching in the unit of whole numbers at the primary education seventh grade mathematics course. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, *24*(2), 25–41.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, *30*(3), 607–610. https://doi.org/10.1177/001316447003000308

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in school children*. University of Chicago Press.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*(3), 151–160. https://doi.org/10.1007/BF02288391

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575.

Leikin, R., Boriskovsky, M., Ovodenko, R., & Miskin, M. (2025). Problem posing or mathematical modeling? The process of expert instructional design. *ZDM–Mathematics Education*, *57*(2), 333–350. https://doi.org/10.1007/s11858-025-01668-1

Lesh, R. A., & Doerr, H. M. (2003). *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching*. Routledge.

Leung, K. C., Leung, F. K., & Zuo, H. (2014). A study of the alignment of learning targets and assessment to generic skills in the new senior secondary mathematics curriculum in Hong Kong. *Studies in Educational Evaluation*, *43*, 115–132. https://doi.org/10.1016/j.stueduc.2014.09.002

Levenberg, İ., & Shaham, C. (2014). Formulation of word problems in geometry by gifted pupils. *Journal for the Education of Gifted Young Scientists*, *2*(2), 28–40.

Lok, B., McNaught, C., & Young, K. (2016). Criterion-referenced and norm-referenced assessments: Compatibility and complementarity. *Assessment & Evaluation in Higher Education*, *41*(3), 450–465. https://doi.org/10.1080/02602938.2015.1022136

Luna, Y. M. B., & Díaz, Y. M. (2023). La habilidad formular problemas matemáticos: Consideraciones para su desarrollo. *Revista Conrado*, *19*(95), 246–253. https://conrado.ucf.edu.cu/index.php/conrado/article/view/3419

Maass, K., Geiger, V., Ariza, M. R., & Goos, M. (2019). The role of mathematics in interdisciplinary STEM education. *ZDM Mathematics Education*, *51*, 869–884. https://doi.org/10.1007/s11858-019-01100-5

Mason, J. (1988). Modelling: What do we really want pupils to learn? In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 201–215). Hodder & Stoughton.

Mason, J. (2005). *Developing thinking in algebra*. Sage.

Mason, J., Burton, L., & Stacey, K. (2010). *Thinking mathematically* (2nd ed.). Pearson.

McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry* (7th ed.). Pearson.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Ministry of National Education (MoNE). (2018). Ortaöğretim matematik dersi öğretim programı (9, 10, 11 ve 12. sınıflar) [High school mathematics curriculum (Grades 9, 10, 11 and 12)]. https://mufredat.meb.gov.tr/Dosyalar/201821102727101-OGM%20MATEMAT%C4%B0K%20PRG%2020.01.2018.pdf

Ministry of National Education (MoNE). (2023). *PISA 2022 Türkiye raporu* [PISA 2022 Türkiye report]. MoNE. https://pisa.meb.gov.tr/meb_iys_dosyalar/2024_03/21120745_26152640_pisa2022_rapor.pdf

Ministry of National Education (MoNE). (2024). *National education statistics formal education 2023/24*. MoNE. https://sgb.meb.gov.tr/meb_iys_dosyalar/2024_10/11230736_meb_istatistikleri_orgun_egitim_2023_2024.pdf

National Council of Teachers of Mathematics (NCTM), (1989). *Curriculum and evaluation standards for school mathematics*. The National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. The National Council of Teachers of Mathematics.

Ng, O. L., & Cui, Z. (2021). Examining primary students' mathematical problem-solving in a programming context: Towards computationally enhanced mathematics education. *ZDM–Mathematics Education*, *53*(4), 847–860. https://doi.org/10.1007/s11858-020-01200-7

Niss, M., & Højgaard, T. (2019). Mathematical competencies revisited. *Educational Studies in Mathematics*, *102*, 9–28. https://doi.org/10.1007/s10649-019-09903-9

Ntumi, S., Agbenyo, S., & Bulala, T. (2023). Estimating the psychometric properties ("item difficulty, discrimination and reliability indices") of test items using Kuder–Richardson approach (KR-20). *Shanlax International Journal of Education*, *11*(3), 18–28. https://doi.org/10.34293/education.v11i3.6081

Organisation for Economic Co-Operation and Development (OECD). (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264190511-en

Organisation for Economic Co-Operation and Development (OECD). (2014). *PISA 2012 technical report*. PISA, OECD Publishing. https://doi.org/10.1787/6341a959-en

Organisation for Economic Co-Operation and Development (OECD). (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. PISA, OECD Publishing. https://doi.org/10.1787/53f23881-en

Özgen, K. (2019). Problem-posing skills for mathematical literacy: The sample of teachers and pre-service teachers. *Eurasian Journal of Educational Research*, *19*(84), 179–212. https://doi.org/10.14689/ejer.2019.84.9

Pérez, H. D. G., Ortega, M. V., & Araque, F. Y. V. (2017). Intervention research and multimethod approach in Human Sciences and mathematical education. *Logos Ciencia & Tecnologia*, *9*(2), 84–95. https://doi.org/10.22335/rlct.v9i2.458

Piaget, J. (1952). *The origins of intelligence in children* (M. Cook, Trans.). W.W. Norton & Co. https://doi.org/10.1037/11494-000

Polya, G. (1945). *How to solve it*. Princeton University Press.

Preciado-Babb, A. P., Saar, C., Marcotte, C., Brandon, J., & Friesen, S. (2013). Using mobile technology for fostering intellectual engagement. *International Journal of Interactive Mobile Technologies (IJIM)*, *7*(3), 46–53. https://doi.org/10.3991/ijim.v7i3.2888

Ramdhani, S., Suryadi, D., & Prabawanto, S. (2019, September). *Some difficulties in making generalization faced by students: A phenomenology study on mathematics learning in Islamic Boarding Schools* [Symposium]. Proceedings of the 1st World Symposium on Software Engineering, pp. 130–134. https://doi.org/10.1145/3362125.3362140

Santos, R., Santiago, A., & Cruz, C. (2024). Problem posing and problem solving in primary school: Opportunities for the development of different literacies. *Education Sciences*, *14*(1), 97. https://doi.org/10.3390/educsci14010097

Santos-Trigo, M. (2024). Problem solving in mathematics education: Tracing its foundations and current research-practice trends. *ZDM–Mathematics Education*, *56*, 211–222. https://doi.org/10.1007/s11858-024-01578-8

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.

Schoenfeld, A. H. (1992). Learning to think mathematically. Problem solving, metacognition and sense-making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). Macmillan.

Stillman, G., Galbraith, P., Brown, J., & Edwards, I. (2007). A framework for success in implementing mathematical modelling in the secondary classroom. *Mathematics: Essential Research, Essential Practice*, *2*(1), 688–697. https://files.eric.ed.gov/fulltext/ED503746.pdf#page = 691

Suna, H. E., Tanberkan, H., & Özer, M. (2020). Changes in literacy of students in Turkey by years and school types: Performance of students in PISA applications. *Journal of Measurement and Evaluation in Education and Psychology*, *11*(1), 76–97. https://doi.org/10.21031/epod.702191

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Taherdoost, H. (2017). Determining sample size; How to calculate survey sample size. *International Journal of Economics and Management Systems*, *2*, 237–239. https://ssrn.com/abstract = 3224205

Tall, D. (2013). *How humans learn to think mathematically: Exploring the three worlds of mathematics*. Cambridge University Press.

Tasarib, A., Rosli, R., & Rambely, A. S. (2025). Impacts and challenges of mathematical modelling activities on students' learning development: A systematic literature review. *Eurasia Journal of Mathematics, Science and Technology Education*, *21*(5), em2641. https://doi.org/10.29333/ejmste/16398

Temel, H., & Altun, M. (2022). The effect of problem-solving strategies education on developing mathematical literacy. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, *16*(2), 275–303. https://doi.org/10.17522/balikesirnef.1164584

Tesfamicael, S. A., & Enge, O. (2024). Revitalizing sustainability in mathematics education: The case of the new Norwegian curriculum. *Education Sciences*, *14*(2), 174. https://doi.org/10.3390/educsci14020174

Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education*. Macmillan.

Toh, T. L., Santos-Trigo, M., Chua, P. H., Abdullah, N. A., & Zhang, D. (Eds.). (2023). *Problem posing and problem solving in mathematics education: International research and practice trends*. Springer Nature Singapore. https://doi.org/10.1007/978-981-99-7205-0.

Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education*, *91*(9), 1426–1431. https://doi.org/10.1021/ed500076x

Treffers, A. (2012). *Three dimensions: A model of goal and theory description in mathematics instruction—The Wiskobas Project* (Vol. 3). Springer Science & Business Media.

Tuna, A., Biber, A. Ç., & Yurt, N. (2013). Mathematical modeling skills of prospective mathematics teachers. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, *33*(1), 129–146.

Ufer, S., Heinze, A., Kuntze, S., & Rudolph-Albert, F. (2009). Proving and substantiating in mathematics education. *Journal Fur Mathematik-Didaktik*, *30*(1), 30–54.

Ülger, T. K., Bozkurt, I., & Altun, M. (2020). Thematic analysis of articles focusing on mathematical literacy in mathematics teaching-learning process. *Education and Science*, *45*(201), 1–37. http://dx.doi.org/10.15390/EB.2020.8028

Van Den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational studies in Mathematics*, *54*, 9–35. https://doi.org/10.1023/B:EDUC.0000005212.03219.dc

Voskoglou, M. G. (2006). The use of mathematical modelling as a tool for learning mathematics. *Quaderni di Ricerca in Didattica*, *16*(1), 53–60.

Wallen, N. E., & Fraenkel, J. R. (2013). *Educational research: A guide to the process*. Routledge.

Walsh, W. B., & Betz, N. E. (1995). *Tests and assessment*. Prentice-Hall.

Wells, C. S., & Hintze, J. M. (2007). Dealing with assumptions underlying statistical tests. *Psychology in the Schools*, *44*(5), 495–502. https://doi.org/10.1002/pits.20241

What Works Clearinghouse (WWC). (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/Handbooks.

Wildani, J. (2020). *The analysis of students' difficulties in solving PISA mathematics problems* [Conference session]. Proceedings of the International Conference on Mathematics and Islam, pp. 246–252.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

Yang, K. L., & Lin, F. L. (2015). The effects of PISA in Taiwan: Contemporary assessment reform. In K. Stacey & R. Tumer (Eds.), *Assessing mathematical literacy: The PISA experience* (pp. 261–273). Springer.

Yu, L., & Li, Y. (2022). A study of practical drawing skills and knowledge transferable skills of children based on STEAM education. *Frontiers in Psychology*, *13*, 1001521. https://doi.org/10.3389/fpsyg.2022.1001521

Zulkardi, Z. (2000). *Realistic mathematics education theory meets web technology* [Conference session]. Proceedings of 10th National Conference of Mathematics. Bandung Institute Technology, Bandung, Indonesia.