The Use of Prompt Engineering in Creating Mathematical Modelling Activities with Artificial Intelligence Tool ChatGPT[1]

Zeynep Gül Dertli
Hacettepe University, Turkey, *zyndrtl@gmail.com*

Bahadır Yıldız
Assoc. Prof. Dr., Hacettepe University, Turkey, *bahadir@bahadiryildiz.net*

Mathematical modelling and modelling activities are important for making sense of mathematical concepts in different extracurricular and daily life contexts. However, teachers may have difficulties in designing these activities in a way to establish meaningful relationships with real life, in accordance with the modeling process and the objectives of the lesson. With a good understanding of how to interact with it, AI tools can be effective tools for teachers in situations where they face difficulties or inadequacies. The aim of the research is to reveal how ChatGPT, an artificial intelligence tool, designs mathematical modelling activities when prompt engineering is applied. ChatGPT has been utilized as a tool for designing mathematical modelling activities at the middle school level. Version 3.5 of ChatGPT has been utilized as the tool for data collection. As part of research, ChatGPT developed mathematical modelling activities using four distinct input formats. Feedback indicated that a gradual increase in performance could be achieved through careful structuring of the inputs. To this end, researchers elaborated on the desired outcome's format, provided examples, and emphasized boundaries as needed. The activities generated by ChatGPT, addressed and examined in terms of appropriation with modelling activity design principles. The research revealed ChatGPT's excellence in designing activities aligning with reality and modelling generalization principles. However, it was unsatisfactory in creating activities aligning with model building, self-evaluation, and model documentation principles. The results indicate that successful prompts for initialization include defining a structure to follow, assigning roles, and using questions for best results. On the other hand, successful prompts for structuring involve defining clear goals, providing good examples, and emphasizing included and excluded information.

Keywords: artificial intelligence, mathematical modelling, model eliciting activities, prompt engineering, ChatGPT

INTRODUCTION

The structure of mathematics education is open to continuous development and thrives on innovation for solving complex problems. In the 21st century, its objective is to offer various learning methods and environments to enable students to mathematically interpret and analyze the situations that they might encounter in their everyday life. Mathematical modeling is one of the ways to provide students with competencies such as defining, making sense of, adapting to, and finding solutions for complex systems that they will encounter in their daily life due to its potential to be the basis of skills such as

---

[1] This study was partially presented as an oral presentation at the International Education Congress (EDUCongress 2023- https://educongress.org/).

critical and analytical thinking and problem solving (Anhalt & Cortez, 2016). Mathematical modeling involves creating problems that stem from real-life scenarios, mathematically formulating these problems to understand them, and utilizing multiple problem-solving cycles to obtain solutions (Lesh & Doerr, 2003). According to Luczak and Earvin (2023), mathematical modeling is an interdisciplinary approach that enhances learning by applying mathematics to real-world problems, is underutilized in contemporary classrooms, and is a unique and powerful part of mathematics. Models are combinations of elements, operations, relationships and rules used to describe, explain, or predict the behavior of a system (Lesh & Doerr, 2003). The process of creating a model begins with a need. The modeling process is a cycle of interpretations and explanations, each step of which can be applied and reconstructed (Doerr & English, 2003).

During the mathematical modeling process, a real-life situation is taken into consideration. This situation is simplified or restructured to make it appropriate for mathematical processing. The structured situation is then translated into a mathematical model. Besides displaying the real situation, this model is also prepared for mathematical operations. The mathematical model undergoes interpretation, suitability checks and appropriate presentations (Blum & Leiss, 2005). According to the Programme for International Student Assessment (PISA) mathematics framework, these processes are also important elements of mathematical reasoning (The Organisation for Economic Co-operation and Development [OECD, 2013]; OECD, 2018). Throughout the cycle, students formulate, operationalize, and evaluate appropriate strategies for using mathematics to solve real-world problems. Through these processes, students engage in mathematical reasoning. Figure 1 illustrates the mathematical modeling process centered on reasoning.
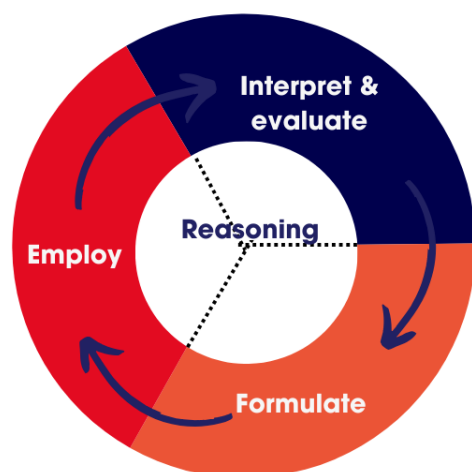

Figure 1
The mathematical modeling cycle of the PISA 2021 framework (OECD, 2018)

Not only does the modeling approach aim to solve problems, but it also aims to create generalizable and adaptable systems that explain relationships (Doerr & English, 2003). The importance of modeling activities using this approach during the process of learning through mathematical modeling has been emphasized in the academic literature (Kaiser & Sriraman, 2006; Lesh & Doerr, 2003).

Mathematical Modelling Activities

Mathematical Modeling Activities are problem-solving process in which students make sense of complex situations, develop their unique mathematical structures and use specific instructional design principles (Kaiser & Sriraman, 2006). It is crucial to characterize a Mathematical Modelling Activity

to ensure intended learning outcomes (Chamberlin & Moon, 2005; Niss, 2010). There are different perspectives on how mathematical modeling and modeling activities should be implemented (Abassian et al., 2020). Lesh et al. (2000) outlined six principles for designing mathematical modeling tasks, which consist of representing reality, constructing models, self-evaluation, documenting the model, generalizing the model, and creating an effective prototype.

The reality principle dictates that the problem situation provided for the modeling activity must be suitable for student's individual knowledge and experiences in their daily life. In accordance with this principle, the modeling activity should be introduced in a practical context and constructed in such a way that students can understand the problem through its connection to their own life (Diefes-Dux et al., 2004; Lesh et al., 2000). Tasks that are familiar to students, such as arranging teams, deciding on the optimal route, and creating shopping lists, can stimulate contemplation of the links between real-life situations and mathematics. By contrast, conventional problems necessitate adherence to procedures (Biccard & Wessels, 2017; Doerr & English, 2003).

The principle of model building is the necessity of creating a model that will explain the goal and possible solution processes of a complex problem situation given in the activity. A typical modeling activity should first reveal the ways of thinking of the students who will participate in the activity (Biccard & Wessels, 2017; Lesh et al., 2000). Therefore, it is important that mathematical modeling activities emphasize students' ways of thinking and create the need for them to create, interpret and explain their own structural systems (Chamberlin & Moon, 2005; Lesh et al., 2000).

The principle of self-assessment postulates that the activity must incorporate standards that students may use to establish and examine their models devoid of any instructor direction (Diefes-Dux et al., 2004). In line with this principle, the activity necessitates having data for gauging the suitability of the solution, the objective of the task must be unambiguous, and pupils ought to have the capacity to decide the course of action necessary to reach their goal (Lesh et al., 2000).

The principle of model documentation requires students to produce a document explaining their problem-solving process (Chen et al., 2022; Doerr & English, 2003). A mathematical modeling task should prompt students to explain how they developed their answer and evaluated the situation in their own words, not just answer the question (Lesh et al., 2000).

The principle of model generalization dictates that the developed model should be adaptable, reusable and shareable as may be required (Lesh et al., 2000). Furthermore, the mathematical model created must not only be unique to the specified situation in the activity, but also valid when applied in comparable situations or utilizing alternative data (Lesh et al., 2000).

The principle of an effective prototype is to prioritize simplicity, comprehensibility, mathematical significance, and potential for future use. The prototype developed by students must be concise, memorable, and applicable to future problem-solving situations (Chen et al., 2022; Lesh et al., 2000).

When a mathematical modeling activity conforms to each of these principles, the mathematical modeling process with students can be carried out more consciously. Students' mathematical thinking and reasoning skills are developed and it becomes possible for them to apply these skills to real-world problems. Teachers who will design mathematical modeling activities to be used in their lessons should take these six principles into consideration (Dede et.al, 2017). However, designing activities consistent with these six principles and encompassing all necessary features of mathematical modeling activities requires dedication and time (Urhan & Dost, 2018). Chamberlin and Moon (2005) affirm that the creation of mathematical modeling activities is possible for anyone, but success requires at least some training on how to create such activities. Teachers need competence in modeling in order to design and conduct their lessons with a focus on mathematical modeling activities, and they need experience outside of school in order for the activities to be realistic (Frejd,

2020). Moreover, the lack of high-quality, relevant examples makes the design of these activities difficult for teachers (Lo et al., 2022). In certain circumstances, it is acknowledged that artificial intelligence can serve as a teaching assistant (Savelka et al., 2023). Due to its noteworthy advancements in recent times, AI can enhance the efficiency of the educational process by aiding educators in tackling complicated issues and making informed decisions (Dwivedi et al., 2023).

Artificial Intelligence and ChatGPT

Artificial intelligence (AI) is the technology that enables a machine to mimic and perform human-like behaviors such as learning, decision-making, problem solving and prediction (Kennedy & Wanless, 2022). Artificial intelligence (AI) is designed to develop algorithms that equip computers with cognitive abilities to undertake tasks that require human intelligence (Abbass, 2019; Coppin, 2004). The progress in large language models (LLMs) has further integrated AI tools into human life, directing research towards new frontiers. Large language models (LLMs) are AI systems that analyze the relationships between words extracted from man-made textual content and use this data for natural language processing tasks that simulate human everyday speech (Melis et al., 2017; Radford et al., 2018). Subjective evaluations are excluded, and technical term abbreviations are explained upon first use. Authors follow conventional academic section and formatting structures, writing in clear, objective, and value-neutral language. They aim for structure with logical flow and causal connections between sentences and paragraphs, avoiding bias and imprecise terminology. Correct spelling, grammar, and punctuation are essential. LLMs have an increasing capacity and have been instrumental in the rise of chatbots, which allow humans to communicate with computers using their own language (Lokman & Ameedeen, 2019). Chatbots with language model algorithms (LLMs) are efficient instruments that generate results based on production objectives and user input (Kuswaha & Kar, 2021; Lokman & Ameedeen, 2019).

One of the most widely used chatbots is ChatGPT, an artificial intelligence language model created by OpenAI (Cooper, 2023). The freely available version 3.5 of ChatGPT is known for its substantial dataset has 175 billion parameters, making it an incredibly effective language model (OpenAI, 2022). ChatGPT has been noted to possess the ability to conduct complex cognitive tasks, generate genuine texts, and furnish effective feedback on mathematical inquiries (Susnjak, 2022; Wardat et al., 2023). Since the emergence and rise of artificial intelligence, it has been recommended to be used by teachers, pre-service teachers and students. In this context, ChatGPT has become an artificial intelligence tool suitable for use by educators due to its potential, rapid diffusion, easy accessibility and easy usability (Elgohary & Al-Dossary, 2023; Nurhasan et al, 2022). However, to leverage ChatGPT's functionalities for these purposes, it must be provided with precise and well-defined inputs (Wardat et al., 2023). This preparation process, regarded as the "most valuable professional skill of this century," is termed prompt engineering (Henrickson & Meroño-Peñuela, 2023).

Prompt Engineering

The process for generating the most appropriate queries or inputs to utilize artificial intelligence tools efficiently is referred to as prompt engineering (Lo, 2023b; Short & Short, 2023). In essence, prompt engineering involves the guidance and programming of a language model to yield optimal outcomes (Reynolds & McDonell, 2021). Prompt engineering is a repetitive refinement procedure that involves testing prompts, scrutinizing responses, and modifying them based on the outcomes in order to achieve optimal results (Lo, 2023a). Large language models (LLMs) are trained using large amounts of textual data to develop a conversational style that is similar to human communication, while at the same time structuring the rules (Dwivedi et al., 2023; Nichols et al., 2019). However, this can result in responses from ChatGPT being too general and out of context, as these models are sensitive to the commands given (Caruccio et al., 2024). Thus, the creation of inputs that engage with artificial intelligence models that process natural language is a vital aspect of interaction (Wang et al., 2023).

Studies indicate that preparing the inputs, also referred to as "prompts," for interactions with conversational artificial intelligence models and aligning them with the desired outputs has a positive impact on these models' performance (Jiang et al., 2020; Yong et al., 2023). Prompt generation and configuration may be carried out through manual interventions or via mining-based and interpretation-based methods, which scrutinize the bonds between words (Jiang et al., 2020). Initially, it is essential to establish a prompt's aim and context in a precise and concise manner. Clear guidance should be provided through the inclusion of appropriate instructions and parameters to achieve the desired outcome (Eager & Brunton, 2023). Personalizing the prompt to determine the intended tone, style or level of detail of the output, and defining the AI's role can enhance the results obtained (Lo, 2023a; Spasić & Janković, 2023). During the refinement process, the quality of results is enhanced by rephrasing with alternate words that focus on specific phrases or relevant keywords that yield superior outputs (Jiang et al., 2020 ; Liu & Chilton, 2022).

When individuals learn to correctly use relevant commands and prompts within their area of specialization, they can maximize AI's potential as an assistant and effectively utilize it as a tool (Eager & Brunton, 2023; Zheng et al., 2023). Likewise, educators who are aware of prompt engineering principles can use AI tools to create accurate, coherent, and relevant content aligned with students' specific needs and learning objectives (Lo, 2023b). A comparable scenario may apply to the mathematical modeling activities under investigation in this study. Well-designed mathematical modeling activities make connections with non-mathematical disciplines and can reveal the importance of using mathematics with these disciplines (Jung & Brand, 2021). In the context of modeling activities, students mathematically interpret a real-life situation (Brady et al.) Moreover, modeling activities involve more than one solution and encourage different solutions (Beckschulte, 2020). These features of modeling activities trigger higher-level thinking. It is believed that modeling activities developed with carefully prepared inputs can aid educators in guiding their pupils through high-level thinking tasks. When used purposefully, artificial intelligence can compensate for teachers' lack of competence and experience in designing mathematical modeling activities to support student learning and higher-order thinking.

The aim of the research is to reveal how ChatGPT, an artificial intelligence tool, creates mathematical modeling activities when prompt engineering is applied.. The research questions are as follows:

- Can Chatgpt create mathematical modelling activities at middle school level?

- How does ChatGPT create mathematical modelling activities when prompt engineering is applied?

METHOD

This research was conducted qualitative approaches. Qualitative methods are favored when a comprehensive understanding of a subject is needed and interpretation is required (Creswell, 2021). The data collection instrument in this research was ChatGPT, an artificial intelligence chatbot created by OpenAI. The study employed the freely available version 3.5 of ChatGPT. In this research, ChatGPT's feature of being a generative artificial intelligence tool was utilized. Thanks to this feature, ChatGPT was made to design mathematical modeling activities and some steps were followed on how the designed activities could be "better". ChatGPT was instructed to produce modeling tasks using four different input formats. These formats were created with the prompts to (1) use general statements without explaining the context, define clear objectives; (2) define a structure to follow and assign roles, use good examples; (3) specify the framework to be used in the assessment, emphasize included and excluded information; and (4) use the questions needed for best results. The responses received were utilized to systematically enhance ChatGPT's modeling task design performance. During the structuring process, the desired output format was specified, exemplified, and its boundaries were emphasized accordingly. The Turkish language was employed to communicate with

ChatGPT. Lo et al. (2023) recommended using clear, concise, and logically structured prompts that can adapt to improve the performance of artificial intelligence language models. The goal is to produce accurate and reliable results. Bias should be avoided, and precise subject-specific vocabulary should be used when appropriate. As such, researchers should provide general statements in the input and customize it to express the form, purpose, scope, and limitations of the output, as well as any specific requirements. Correct grammar, spelling, and punctuation are also necessary for achieving academic writing quality. The restructuring process with new inputs was gradually conducted until each principle in the evaluation framework was appropriately addressed. The activities, along with their structured versions, were evaluated based on the principles of modeling activity design.

Data Analysis

The data collected were analyzed using content analysis. The modeling activities were created and revised based on various inputs. The framework proposed by Lesh et al. (2000) for designing modeling activities, which includes principles such as reality, model building, self-evaluation, model documentation and model generalization, was taken into consideration. Each activity generated by ChatGPT was examined in terms of its appropriation with these principles. The effective prototype principle was not evaluated as it necessitates a certain period for implementing activities and evaluating outcomes. The activities were categorized as "completely appropriate", "partially appropriate" and "not appropriate" according to their fulfillment to meet each of the five principles with reference to similar studies in the literature. If the activity presented a scenario that highlights a problem commonly faced by students in their daily life, and if it clearly demonstrated the need for the students' assistance, it was considered appropriate for the reality principle. Even if a situation that can be encountered in real life is given in the activity, if this situation is not presented in a meaningful context for students, it partially appropriate with the reality principle. Activities requiring tasks such as "creating a model," "writing an equation," or "writing a mathematical expression" were fully compliant with the modeling principle. In contrast, situations that necessitate creating a model to solve the problem were deemed partially appropriate in relation to the modeling principle, despite not being explicitly mentioned in the activity. Situations where the solution of the problem does not require the creation of a mathematical model, but can only be solved using operations, is not appropriate with the principle of model building. If the activity employed phrases such as "debate with your group," "clarify your approach," or "scrutinize the precision of your answer," it was assessed as fully appropriate for the principle of self-assessment. Apart from this, activities were deemed partially appropriate for the principle of self-evaluation if the data were clear and understandable to the students and the problem was appropriate for their level. Activities that required students to document and elaborate the models they formulated, such as "explain to the customer," "submit to the authorities," and "write a letter to clarify...," were fully compatible with the principle of model documentation. If the activity does not include any expression for students to express their solutions through a written document, the model is not appropriate in terms of the model documentation principle. If the model created for the given problem situation can serve the same purpose for different people, it is fully compatible with the principle of model generalization. If a model is only functional in a specific situation or for specific people, if it is not suitable for use by different users, the model is not appropriate according to the principle of generalization (Deniz & Akgün, 2016; Lesh et al., 2000; Sezen Yüksel et al., 2019). The researchers individually categorized the activities based on their compliance with the principles and found a harmonious relationship between their analyses.

FINDINGS

This section the activities and assessments generated by ChatGPT in response to the prompts of defining clear objectives through contextualization, assigning roles by establishing a structured plan with good examples, specifying the assessment framework with information that is included and excluded, and utilizing necessary questions to obtain optimal results.

Activity 1:

The initial task assigned to ChatGPT was to develop a mathematical modeling activity that meets the needs of middle school students. The activity designated as A1.I1. The language is free of ornamental or emotional elements, with no contractions or colloquialisms. The structure of the activity is easy to follow with causal connections between statements. The vocabularies used are technical and precise, avoiding the error of using similar non-technical terms. The text is free of grammatical and spelling errors, employs standard sentence structures, and adheres to common academic sections. The activity maintains a formal register while avoiding bias, and technical term abbreviations are explained for clarity. All citations are consistent with prescribed guidelines, and quotes are clearly marked. The initial task assigned to ChatGPT was to develop a mathematical modeling activity that meets the needs of middle school students. The initial task assigned to ChatGPT was to develop a mathematical modeling activity that meets the needs of middle school students. It was designed with a focus on clear and concise language, logical progression, and objective presentation of information. A1 I1., is illustrated in Figure 2.

Can you design a mathematical modeling activity suitable for middle school students?

You and your friends decide to get together for the weekend and plan to organize a pizza party. However, each of you has to work with a certain budget. Your total budget is 100 TL and you decide to make the best use of it by ordering pizzas.

On the menu of the pizza shop there are different sizes of pizzas and prices of different toppings. Each of you has to think about how many people you are going to order pizza for and which pizza sizes and toppings to choose. The aim is to create a pizza order that appeals to the largest number of people and provides the most pleasure within the given budget.
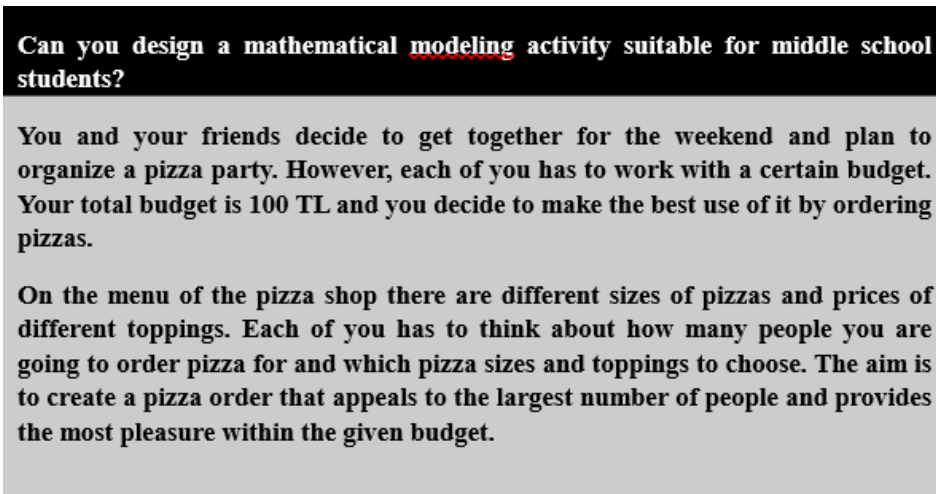
Figure 2
A1.I1 designed by ChatGPT (Translated to English)

The activity created with this input (A1. I1.) exhibited inadequacies in model construction, self-assessment, and adherence to model documentation principles. Students were presented with a scenario of selecting the most suitable pizza party option as a problem situation. This problem was contextualized to simulate a real-life scenario for students, which aligns with the reality principle. The students were tasked with designing an order that maximizes the budget and appeals to the greatest number of people. However, it was not explicitly stated that a mathematical model should be created to express this order. As a result, the activity was deemed only partially compliant with the model-building principle. In the given scenario, the pizza shop's menu comprises various pizza options with diverse sizes and ingredient prices, although ChatGPT did not produce this data. Additionally, there is no provision for students to assess their answers. Nonetheless, considering the problem's suitability for middle school students, it can be categorized as partially appropriate for self-evaluation. The budgeting and pricing procedures given in the activity are suitable for middle school level. Since students will be working with a familiar context and data, they will be able to analyze the problem and thus objectively evaluate their own performance. The activity fails to comply with the model documentation principle since it lacks written documentation for students to articulate their solutions.

However, as the created model can be applied to comparable situations, the activity is fully compliant with the model generalization principle.

To improve the activity in terms of the principles in which it was inadequate, explicit goals were set for these principles and the activity was asked to be rewritten. For the modeling principle, the input was "Can you rewrite this activity in a way that requires students to create a model?". The activity created by ChatGPT according to this input named A1.I2. is given in Figure 3.

Can you rewrite this activity to require students to come up with a model?

Of course, a mathematical modeling activity that requires students to create a model can be designed as follows:
Activity Name: Pizza Party Cost Modeling and Decision-Making
Target Age Group: Middle school students (6th grade and above)
Objective: To develop students' mathematical modeling skills and teach them to create a mathematical model by analyzing a real-world problem.
Encourage students to create their models using a mathematical expression or equation. For example, to calculate the cost of ordering a pizza, they could use an equation such as
"Total cost = (Pizza unit price) x (How many pizzas we order)".
Ask each student to create their model and find the best pizza order within a given budget. Using their modeling, students should identify the pizza orders that appeal to the largest number of people and make the best use of the budget.
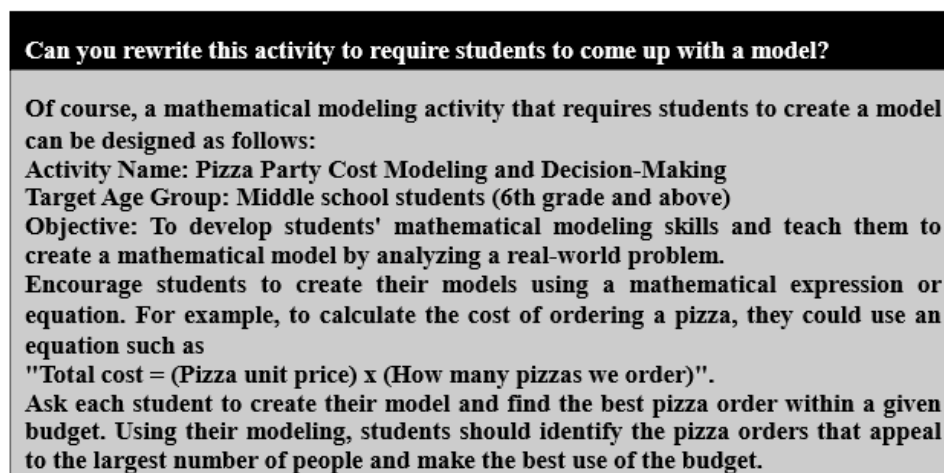
Figure 3
A1.I2 designed by ChatGPT

It was found that ChatGPT included "an equation to calculate the cost of a pizza order" and "developing a model using a mathematical expression or equation" in the activity presented. Therefore, this activity fully complies with the model-building principle. Furthermore, by specifying "presenting the model as a report" in the activity, the model also adheres to the documentation principle.

Similarly, following the self-evaluation principle, the task was revised based on the prompt "Can you modify this activity so that students can assess their solutions independently of any teacher guidance?" Accordingly, ChatGPT incorporated the phrase "conduct self-evaluations and exchange evaluations with classmates" into the task. Similarly, following the self-evaluation principle, the task was revised based on the prompt "Can you modify this activity so that students can assess their solutions independently of any teacher guidance?" Accordingly, ChatGPT incorporated the phrase "conduct self-evaluations and exchange evaluations with classmates" into the task. Consequently, the task now fully meets the standards for self-evaluation. The final version of Activity 1 called A1.I3., which was structured with the new inputs, fully complies with the characteristics of a mathematical modeling activity. The findings obtained from the analysis of Activity 1 created by ChatGPT are given in Table 1.

Table 1

Findings from the analysis of activity 1 created by ChatGPT

Principles	Fully compliant	Partially compliant	Not compliant
Reality	A1.I1., A1.I2., A1.I3.		
Model creation	A1.I2., A1.I3.	A1.I1.	
Self-assessment	A1.I3.	A1.I1., A1.I2.	
Model documentation	A1.I2., A1.I3.		A1.I1.
Model generalization	A1.I1., A1.I2., A1.I3.		

Activity 2

This activity was designed by defini

ng a structure and role to achieve academic writing quality. The focus was on objectivity, comprehensibility, logical structure, conventional structure, clear, objective language, format, formal register, structure, balanced approach, precise word choice, and grammatical correctness. The activity aimed to answer the following questions - what is mathematical modeling? What are its components and activities? How can mathematical modeling activities be evaluated? After completing the questionnaire, ChatGPT received the task of creating a mathematical modeling activity for middle school students in the role of a mathematics teacher and education researcher. The activity generated by ChatGPT based on input A2. I1. is displayed in Figure 4.

> **Answer as a math teacher and math education researcher. Can you design a mathematical modeling activity at the middle school level?**
>
> Of course, I can design a mathematical model building activity at the middle school level.
> Below is an activity I designed as a math teacher and educational researcher:
> **Activity Name: Modeling Air Temperature**
> **Objective:** This activity aims to give students the ability to model air temperature mathematically.
> **Objectives:**
> - To improve students' data collection and analysis skills.
> - To teach students the mathematical modeling process.
> - To develop students' ability to make predictions based on their model.
> **Materials:**
> - Air temperature data (monthly temperature averages or daily temperature data)
> - Graph paper
> - Calculator
> - Computer or tablet (optional)
> **Activity Stages:**
> **Step 1: Problem Identification (Class Discussion - 15 Minutes)**
> Ask students a question about how air temperature changes through the seasons.
> Sample question: "How does the air temperature change and what are the reasons for these changes?"

Figure 4

A2.I1 designed by ChatGPT

Activity A2. I1 involves predicting changes in air temperature. There was a real-world scenario in the exam, but there was a lack of clear and meaningful context for students. Thus, the activity was only partially appropriate for real-world application. The activity was considered fully appropriate for model building because it included the objective of "creating a mathematical model to predict changes in air temperature." However, there was no guidance for students to evaluate the solution processes. The activity was assessed as partially appropriate for the principle of self-evaluation, as it was designed for the middle school level. Since the activity includes the instruction that each group should prepare and present a document that communicates the model they created and the results of their predictions, the model is fully compliant with the documentation principle. Additionally, since the model created can be applied to similar situations in the future, the activity has been deemed fully compliant with the principle of model generalization.

In order to improve the efficiency in terms of the reality principle, a new input was given and a good example was used as a reference for the answer. Specifically, the task requires that students model air temperature. However, there is insufficient information regarding the purpose and need for student assistance. Lesh and Doerr's (2003) modeling activity, the Big Foot problem, is an accepted exemplar in the literature. Using this example as a reference, ChatGPT included a problem scenario that is completely appropriate to the principle of reality during the activity reconstruction. This scenario involves helping a farmer to harvest at the correct time. The resulting activity labeled A2.I2., can be seen in Figure 5.

In this activity, we ask students to model the air temperature, but there is no information about what they need to do this for, or what they need the students' help for. Can you explain this to the students in an open-ended context? For example, "The police found footprints at a murder scene. The police want to make a guess about the height of the criminal based on the footprint seen below. To help the police, develop a mathematical formula to find the height of the person from the footprint."

Activity Name: Air Temperature Modeling - Agriculture and Harvest Forecasting
Objective: This activity aims to teach students to understand air temperature through a mathematical modeling process and to use this knowledge in agriculture to predict the harvest.
Scenario:
Students will meet Mr./Mrs. Smith, who owns a farm. Mr./Mrs. Smith grows different crops on his farm and he knows that air temperature is an important factor when harvesting these crops.

Mr./Mrs. Smith tells the students the following: "In order to harvest at the right time, we realized that changes in air temperature have a significant impact on plant growth. For example, some plants like warm weather while others prefer cooler conditions. Therefore, can you help us build a prediction model based on air temperature?"

Figure 5
A2.I2 designed by ChatGPT

To align with the principle of self-evaluation, students must assess their findings independently without teacher guidance. To ensure compatibility with the principle of self-evaluation, students should present their findings in an unbiased, objective manner. This can be accomplished through a

clear and straightforward presentation of data within the activity. The first and second iterations of Activity 2 (A2.I1., A2.I2.) included directives that guided students through each step. Furthermore, these versions lack a dataset that students may utilize. As a result, in the following stage, ChatGPT was tasked with redoing the activity while highlighting the included and excluded information. The instructions given to students in the activity were omitted as per the input, and the dataset necessary for the solution was included. The activity created by ChatGPT according to this input, called A2.I3, is shown in Figure 6.

I want the students participating in this activity to think and decide for themselves what they need to do, so could you redesign this activity without explicitly stating what they need to do, and by creating and including the air temperature data and plant data yourself?

Of course, I can redesign the activity to give students more autonomy. Here is a free example scenario where students can create and use their own data and need to model air temperature:

Activity Name: Air Temperature Modeling - Plant Growth Prediction

Objective: This activity aims to teach students to understand air temperature through a mathematical modeling process and use this knowledge to predict plant growth and harvest.
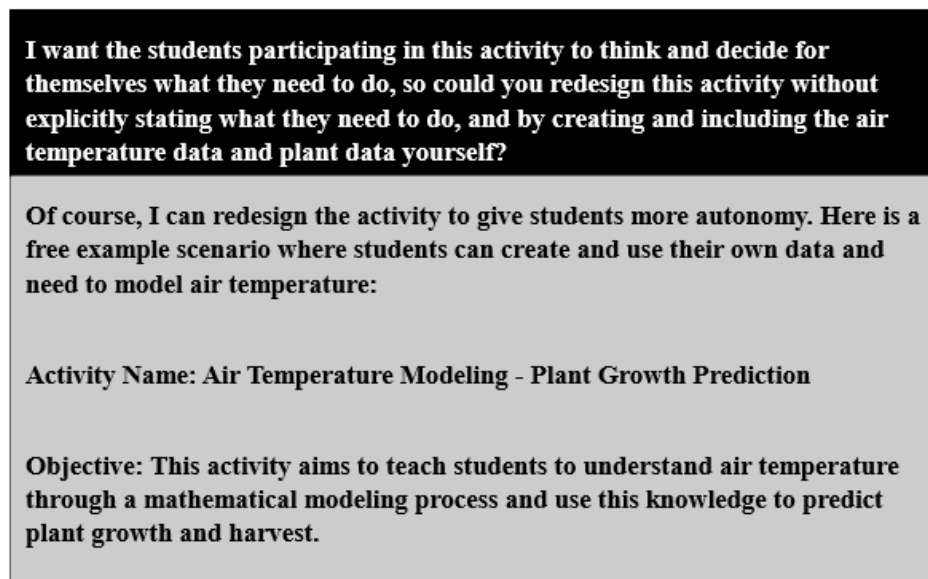
Figure 6
A2.I3 designed by ChatGPT

The final version of Activity 2 (A2.I3.), which was structured with new inputs, fully conforms to the characteristics of a mathematical modeling activity. The findings obtained from the analysis of Activity 2 created by ChatGPT are given in Table 2.

Table 2
Findings from the analysis of activity 2 created by ChatGPT

Principles	Fully compliant	Partially compliant	Not compliant
Reality	A2.I2., A2.I3.	A2.I1.	
Model creation	A2.I1., A2.I2., A2.I3		
Self-assessment	A2.I3	A2.I1., A2.I2.	
Model documentation	A2.I1., A2.I2., A2.I3.		
Model generalization	A2.I1., A2.I2., A2.I3		

Activity 3

This activity was developed by outlining the framework for the assessment based on the input received by ChatGPT, which was to create a mathematical modeling activity for middle school students that adhere to all of Lesh et al.'s (2000) principles of reality, model-building, model documentation, model generalization, and self-evaluation. ChatGPT designed an activity that involves a mathematical evaluation of whether library cards are fair. The activity created by ChatGPT according to this input, named A3. I1., is given in Figure 7.
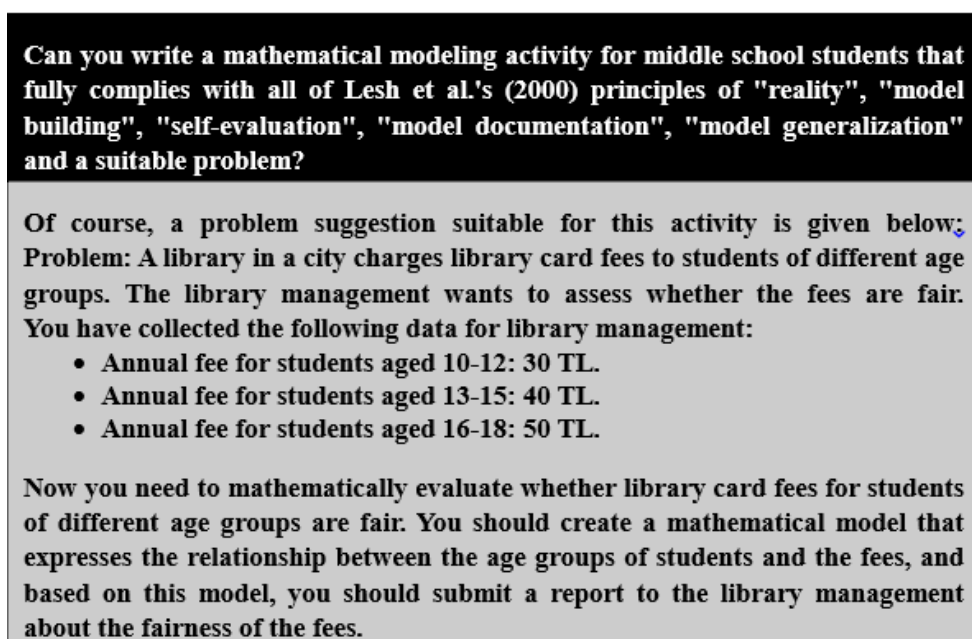
Can you write a mathematical modeling activity for middle school students that fully complies with all of Lesh et al.'s (2000) principles of "reality", "model building", "self-evaluation", "model documentation", "model generalization" and a suitable problem?

Of course, a problem suggestion suitable for this activity is given below:
Problem: A library in a city charges library card fees to students of different age groups. The library management wants to assess whether the fees are fair. You have collected the following data for library management:

- Annual fee for students aged 10-12: 30 TL.
- Annual fee for students aged 13-15: 40 TL.
- Annual fee for students aged 16-18: 50 TL.

Now you need to mathematically evaluate whether library card fees for students of different age groups are fair. You should create a mathematical model that expresses the relationship between the age groups of students and the fees, and based on this model, you should submit a report to the library management about the fairness of the fees.

Figure 7
A3.I1 designed by ChatGPT

The activity fully complies with the principles of reality, model building, model documentation, and model generalization, and partially complies with the principle of self-evaluation. The activity problem is contextualized in a real-life scenario, thus complying entirely with the reality principle. It fully complies with the model-building principle as the problem states the task of creating a mathematical model that expresses the relationship between age groups and wages. Additionally, it adheres to the model documentation principle as it indicates the presentation of a report to the library management based on this model. Since the model created to solve this problem can be applied to comparable situations, the activity aligns with the principle of model generalization. There was no statement provided for the students to assess and validate the accuracy of the solutions. However, as the data used for the solution was lucid and comprehensible, the activity partly adheres to the principle of self-evaluation.

To improve the activity in terms of the self-evaluation principle, it was asked to rewrite the activity by specifying an explicit goal for this principle, similar to the one in Activity 1. ChatGPT reconstructed the activity for this input by adding the statement "evaluate the reliability and shortcomings of the model". The final version of Activity 3 (E3.I2.), which was structured with the new inputs, fully conforms to the characteristics of a mathematical modeling activity. The findings obtained from the analysis of Activity 3 created by ChatGPT are given in Table 3.

Table 3
Findings from the analysis of activity 3 created by ChatGPT

Principles	Fully compliant	Partially compliant	Not compliant
Reality	A3.I1., A3.I2.		
Model creation	A3.I1., A3.I2.		
Self-assessment	A3.I2.	A3.I1.	
Model documentation	A3.I1., A3.I2.		.
Model generalization	A3.I1., A3.I2.		

Activity 4

This activity was created using the input "Ask me at least 10 questions to get the best results." For this purpose, a mathematics teacher and educational researcher asked me to design mathematical modeling activities and problems to use in lessons. This activity was created using the input "Ask me at least 10 questions to get the best results." For this purpose, a mathematics teacher and educational researcher asked me to design mathematical modeling activities and problems to use in lessons. Before starting, there must be clarification about the type of activities needed. For this input, ChatGPT requested detailed information on mathematical skills, topics, and concepts that should be emphasized in the activity, as well as the appropriate level of difficulty for students. Collaboration was also strongly encouraged, and concrete and feasible solutions were required. Researchers were asked to consider the available resources for students when designing the activity. All questions posed by ChatGPT were answered thoroughly. The activity created by ChatGPT, named A4. I1 is presented in Figure 8.

I am a mathematics teacher and educational researcher. I am going to ask you to design mathematical modeling activities and problems to use in my classes. Before we start, I want you to understand exactly what kind of activities I want. Ask me at least 10 questions about the activities so that you can get the best results.

Of course, I can help you design mathematical modeling activities and problems. Here are 10 questions I need to answer to help you:

1. In which math topics would you like to design modeling activities and problems?

2. For which age groups or levels of students do the activities or problems need to be?

3. What mathematical concepts or skills do you want your students to learn or practice from these activities or problems?

4. Should your activities or problems be based on real-world applications or emphasize abstract mathematical concepts?

5. Should I make the activities group work or project-based to encourage student collaboration?

6. Should your activities or problems have concrete results or feasible solutions?

7. What resources or tools do you expect students to use when working on these activities (e.g. calculators, computers, graphing programs)?

8. What should be the level of complexity of the activities or problems? Do you prefer them to be simple, moderate or challenging?

9. What steps do you expect students to follow when solving the problems?

10. Do you want me to emphasize any specific teaching goals or objectives in your activities or problems?

Problem: "Coffee Sales and Business Strategy"

Description:

A group of students running a school canteen want to increase coffee sales and run the canteen more efficiently. They need to use the following information to develop business strategies.

Question:

Which types of coffee are most in demand in the school canteen?

At what times are coffee sales highest and lowest?

What should the price of coffee be so that customers are satisfied and profits are made?

Which ingredients need to be ordered and what is the total cost of these ingredients? How much profit can the business make on a monthly basis?

Based on this information, create and present to the class a business plan to increase coffee sales and make the canteen more profitable. Also, prepare a memo explaining your business plan.

Figure 8
The questions ChatGPT asks and the A4.I1

In this activity, students were given a problem in which they were asked to create a plan to increase the sales of coffee in the school cafeteria and the profit from those sales. The activity was categorized as fully appropriate for the principles of reality, model documentation, and model generalization, and partially appropriate for the principles of model building and self-evaluation. The problem presented in this activity was presented in the context of a scenario that students might encounter in real life, so it was fully consistent with the Reality Principle. The activity requires the creation of a business plan, but it does not specify the need for a mathematical model to explain the solution. As a result, the activity partially complies with the modeling principle. While the instructions indicated presenting the business plan in the classroom, it only partially complied with the self-evaluation principle due to the lack of clarity about the purpose of the presentation. The activity fully adhered to the model documentation principle, as demonstrated by the requirement to draft a memo outlining the business plan. Since the created model for solving this issue can apply to comparable situations, the task is assessed as fully compliant with the principle of model generalization.

To structure the activity problem, ChatGPT was prompted to ask 10 additional questions in response to the second input which stated: "This problem did not entirely meet my expectations. Ask me 10 more questions to structure the problem." ChatGPT then inquired about the possibility of developing a list featuring various coffee types and their corresponding prices, techniques for guiding students in modeling and presenting viable solution strategies, in addition to methods for promoting collaborative evaluations of the generated results. After answering each question in detail, the teacher revised the activity (A4. I2.) with consideration of the responses. The revised activity aligned with the characteristics of a mathematical modeling task. Table 4 presents the findings from the analysis of Activity 4 created by ChatGPT.

Table 4
Findings from the analysis of activity 4 created by ChatGPT

Principles	Fully compliant	Partially compliant	Not compliant
Reality	A4.I1., A4.I2.		
Model creation	A4.I2.	A4.I1.	
Self-assessment	A4.I2.	A4.I1.	
Model documentation	A4.I1., A4.I2.		
Model generalization	A4.I1., A4.I2.		

DISCUSSION

It was investigated how the mathematical modeling activities generated by ChatGPT are in terms of modeling principles and how these activities are when prompt engineering is applied. The findings indicated that ChatGPT, an artificial intelligence chatbot, has the potential to successfully generate mathematical modeling activities. To fully realize and utilize the potential of ChatGPT, effective communication inputs are necessary. As stated by Jiang et al. (2020), when unsuitable prompts are utilized with artificial intelligence language models, their knowledge is constrained and their potential is not fully revealed. Vermeersch (2023) contends that formulating the input format and context to achieve the desired outputs results in precise and pertinent responses.

It was found that ChatGPT was successful in designing modeling activities that comply with the reality principle. ChatGPT was capable of producing modeling activities that adhered to the reality principle irrespective of the input. ChatGPT has included a real-life situation in all of its designed activities. The problem expression in the activity was affected by prompt engineering in relation to this principle. In three activities, it created real-life problems in the context of scenarios that were meaningful for students. In one activity where it did not do this, it compensated when prompt engineering was applied. While ChatGPT does not directly address the self-evaluation principle in

modeling activities, it has the ability to design activities that partially conform to this principle. With the application of demand engineering, suitable self-evaluation expressions were added to the activity. While most of the modeling activities generated by ChatGPT were partially appropriate or inappropriate in terms of complying with the model-building principle, it was determined that they could be made appropriate by including explanatory inputs. When evaluating modeling activities in terms of principles and adequacy, it is worth noting that similar results have been found in assessing activities created by humans in existing literature. Deniz and Akgün (2016) noted in their research that the model-building activities developed by secondary mathematics teachers were partially appropriate for the principle of self-evaluation. Similarly, Urhan and Dost (2018) found that all modeling activities in a ninth-grade math textbook partially aligned with the principle of self-evaluation while most partially aligned with the principles of model building and documentation. Similarly, in a study evaluating mathematical modeling activities developed by pre-service math teachers, Tekin Dede et al. (2017) found similar results. These findings confirm Li et al.'s (2023) perspective that the focus should not be on ChatGPT suggestions surpassing human ones, but rather on their potential to generate better outcomes through some settings and modifications.

ChatGPT stated that in modeling activities, it is necessary to create expressions such as the statement of the problem situation, alternative scenarios, and a data set for the solution. Although ChatGPT did not create these expressions itself, it was able to generate them when provided with the appropriate inputs. ChatGPT also demonstrated the ability to continue a conversation in the previous context. This function enables the generation of prompts without the need for frequent rewriting (Dwivedi et al., 2023; Henrickson & Meroño-Peñuela, 2023). Additionally, Zheng et al. (2023) reported testing and improving the prompts by adding specific details, providing feedback to ChatGPT's outputs, and enhancing overall quality.

When structuring the activities, prompts to define clear objectives, use effective examples, and highlight included and excluded information had a positive impact on the quality of the modeling activities. Correspondingly, Zheng et al. (2023) found that by emphasizing which aspects ChatGPT should ignore and focus on in their inputs, they were able to prevent undesired outcomes and improve the effectiveness of filtering the results.

When compared using the evaluation framework and writing style principles, it is evident that Activity 1, which lacked contextual explanation and relied on general statements, was less successful compared to Activity 3, which provided an evaluation framework as a reference and displayed relatively more effectiveness. Defining a clear structure, assigning roles, and utilizing targeted questioning were identified as effective prompts for achieving optimal outcomes. Correspondingly, in their study on devising a programming lesson plan for preschoolers via ChatGPT, Spasić and Janković (2023) found that role assignment, contextualization, and elaboration prompts resulted in superior results compared to standard prompts lacking context. Wang and colleagues (2023) reported a difference in the results of prompts created in three different formats: simple, detailed, and with examples.

Research indicates that while ChatGPT use in educational contexts may have advantages, it also presents limitations and potential risks (Chassignol et al., 2018; Farrokhnia et al., 2023; Susnjak, 2022). Artificial intelligence tools, like ChatGPT, have the potential to enhance work efficiency by stimulating teacher creativity in developing adaptive teaching materials (Dwivedi et al., 2023; Lee & Yeo, 2022; Xia et al., 2022). However, excessive reliance on ChatGPT may detrimentally impact the quality of education and teacher motivation (Farrokhnia et al., 2023). It's important to strike a balance between utilizing technology and promoting creativity and motivation in the educational process. As observed in this study, ChatGPT is capable of producing results of varying quality depending on the provided request, and its output can be enhanced through improved requests. Additionally, human creativity may still outperform the quality of ChatGPT's content on certain occasions. The study

results offer insights into how to mitigate potential educational threats from ChatGPT by establishing effective communication tactics. Moreover, it demonstrates the feasibility of transforming ChatGPT into an effective educational tool.

CONCLUSION AND SUGGESTIONS

In this study, it was revealed how ChatGPT, a generative artificial intelligence tool, can be used to design mathematical modelling activities at the middle school level. In addition, it was also seen how prompt engineering can improve the created activities in terms of mathematical modelling principles. Even when ChatGPT is asked to design a mathematical modeling activity without any preparation or prompt engineering, it produces results that can generate ideas for teachers. However, when some of the subtleties of asking questions to ChatGPT are known and used, when prompt engineering is applied, the results can be improved in terms of the principles of mathematical modeling activity generation. Of the input formats tested in this research, the combination of "specify the framework to be used in the assessment, emphasize included and excluded information" appeared to yield the best results. However, it should be kept in mind that these results were obtained at the time and under the conditions under which this research was conducted, and that different results may be obtained by future researchers due to the continuous updating of productive artificial intelligence tools such as ChatGPT. Within the scope of the research, different forms of input were also tested and the results were presented as a guide for future researchers. In a different research context, any of these inputs may yield better results. This study has some limitations. The process of designing the mathematical modeling domain, which is the subject of this research, was carried out with version 3.5 of ChatGPT. This process can produce different, more comprehensive results with version 4, which can work on a larger dataset. The results of this study were analyzed within the scope of a selected modelling framework and the activities were not implemented in the classroom. Therefore, the results are limited to the theoretical framework. The learning objectives for the activities created in this study were ignored. The activities are limited to middle school level.

According to the research results;

- ChatGPT can be used as a teacher assistant in math education.

- By communicating efficiently, ChatGPT can help mathematics teachers design content that requires time and effort, such as modeling activities.

- Despite its impressive results, teachers need to try and test AI-generated content before using it in their classrooms, and evaluate its accuracy and appropriateness.

- More accurate and consistent results can be obtained by using artificial intelligence tools by considering prompt engineering and its effects.

For future research;

- Research and guidance for teachers who want to use AI as an auxiliary tool in their lessons should be increased.

- The results of this research can be extended by experimenting with different inputs and prompt patterns.

- The activities designed within the scope of this research can be implemented in the classroom environment and reflections from the implementation process can be shared.

- Mathematical modeling activities can be designed with a different Chatbot or with the paid version 4 of ChatGPT and the results can be compared.

REFERENCES

Abbass, H. A. (2019). Social integration of artificial intelligence: functions, automation allocation logic and human-autonomy trust. *Cognitive Computation, 11(2),* 159-171. https://doi.org/10.1007/s12559-018-9619-0

Abassian, A., Safi, F., Bush, S., & Bostic, J. (2020). Five different perspectives on mathematical modeling in mathematics education. *Investigations in Mathematics Learning*, *12*(1), 53-65. https://doi.org/10.1080/19477503.2019.1595360

Anhalt, C., & Cortez, R. (2016). Developing understanding of mathematical modeling in secondary teacher preparation. *Journal of Mathematics Teacher Education*, 19, 523-545. https://doi.org/10.1007/S10857-015-9309-8.

Beckschulte, C. (2020). Mathematical modelling with a solution plan: An intervention study about the development of grade 9 students' modelling competencies. In: Stillman, G.A., Kaiser, G., Lampen, C.E. (eds) *Mathematical Modelling Education and Sense-making. International Perspectives on the Teaching and Learning of Mathematical Modelling* (pp. 129-138). Springer, Cham.

Biccard, P., & Wessels, D. (2017). Six principles to assess modelling abilities of students working in groups. In: Stillman, G., Blum, W., Kaiser, G. (eds) *Mathematical modelling and applications. International perspectives on the teaching and learning of mathematical modelling.* Springer, Cham. https://doi.org/10.1007/978-3-319-62968-1_49

Blum, W., & Leiss, D. (2005, February). Filling Up "-the problem of independence-preserving teacher interventions in lessons with demanding modelling tasks. In *CERME 4–Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education* (Vol. 1623).

Brady, C., McLean, J. A., Jung, H., Glancy, A. W., & Dominguez, A. (2020). Investigating the complexity of student modelling by coordinating a diversity of research perspectives. In: Stillman, G.A., Kaiser, G., Lampen, C.E. (eds) *Mathematical Modelling Education and Sense-making. International Perspectives on the Teaching and Learning of Mathematical Modelling* (pp. 75-84). Springer, Cham.

Caruccio, L., Cirillo, S., Polese, G., Solimando, G., Sundaramurthy, S., & Tortora, G. (2024). Can ChatGPT provide intelligent diagnoses? A comparative study between predictive models and ChatGPT to define a new medical diagnostic bot. *Expert Systems with Applications, 235,* 121186. https://doi.org/10.1016/j.eswa.2023.121186

Chamberlin, S. A., & Moon, S. M. (2005). Model-eliciting activities as a tool to develop and identify creatively gifted mathematicians. *Journal of Secondary Gifted Education, 17(1),* 37-47. https://doi.org/10.4219/jsge-2005-

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: a narrative overview. *Procedia Computer Science, 136*, 16-24. https://doi.org/10.1016/j.procs.2018.08.233

Chen, G. X., Qu, X. Q., Huang, L. P., Huang, L., Zhou, C., & Qiao, M. Y. (2022). Modeling-eliciting activities in an online engineering course for improving conceptual learning, professional skill, interaction. *IEEE Access*, *10*, 87767-87777. https://doi.org/10.1109/ACCESS.2022.3199690

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology, 32(3),* 444-452. https://doi.org/10.1007/s10956-023-10039-y

Coppin, B. (2004). *Artificial intelligence illuminated*. Jones & Bartlett Learning.

Creswell, J. W. (2021). *Nitel araştırma yöntemleri: Beş yaklaşıma göre nitel araştırma ve araştırma deseni.* (3. Baskıdan Çeviri).

Deniz, D., & Akgün, L. (2016). Ortaöğretim matematik öğretmenlerinin model oluşturma etkinliği tasarım prensiplerine uygun etkinlik tasarlayabilme yeterlikleri. *Karaelmas Eğitim Bilimleri Dergisi, 4(1),* 1-14.

Dede, A.T., Hıdıroğlu, Ç.N., & Güzel, E.B. (2017). Examining of model eliciting activities developed by mathematics student teachers. *Journal on Mathematics Education*, 8(2), 223-242. https://doi.org/10.22342/jme.8.2.3997.223-242

Diefes-Dux, H. A., Moore, T., Zawojewski, J., Imbrie, P. K., & Follman, D. (2004, October). A framework for posing open-ended engineering problems: Model-eliciting activities. *In 34th Annual Frontiers in Education, 2004.* FIE 2004. (pp. F1A-3). IEEE. https://doi.org/10.1109/FIE.2004.1408556

Doerr, H. M., & English, L. D. (2003). A modeling perspective on students' mathematical reasoning about data. *Journal for research in mathematics education, 34(2),* 110-136. https://doi.org/10.2307/30034902

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71,* 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Eager, B., & Brunton, R. (2023). Prompting higher education towards AI-augmented teaching and learning practice. *Journal of University Teaching & Learning Practice, 20(5)*, 02. https://doi.org/10.53761/1.20.5.02

Elgohary, H. K. A., & Al-Dossary, H. K. (2023). The effectiveness of an educational environment based on artificial intelligence techniques using virtual classrooms on training development. *International Journal of Instruction, 15(4),* 1133-1150. https://doi.org/10.29333/iji.2022.15460a

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International, 1-15.* https://doi.org/10.1080/14703297.2023.2195846

Frejd, P. (2020). Teachers' possibilities to generate science capital for modelling. In: Stillman, G.A., Kaiser, G., Lampen, C.E. (eds) *Mathematical Modelling Education and Sense-making. International Perspectives on the Teaching and Learning of Mathematical Modelling* (pp. 39-49). Springer, Cham. https://doi.org/10.1007/978-3-030-37673-4_2

Henrickson, L., & Meroño-Peñuela, A. (2023). Prompting meaning: a hermeneutic approach to optimising prompt engineering with ChatGPT. *AI & SOCIETY, 1-16.* https://doi.org/10.1007/s00146-023-01752-8

Jiang, Z., Xu, F. F., Araki, J., & Neubig, G. (2020). How can we know what language models know? *Transactions of the Association for Computational Linguistics, 8,* 423–438. https://doi.org/10.1162/tacl_a_00324

Jung, H., & Brand, S. (2021). Student actions for social justice-oriented mathematical tasks. *Mathematics Teacher: Learning and Teaching PK–12, 114(5), 378–385.*

Kaiser, G., & Sriraman, B. (2006). A global survey of international perspectives on modelling in mathematics education. *ZDM, 38,* 302-310. https://doi.org/10.1007/BF02652813

Kennedy, H., & Wanless, L. (2022). Artificial intelligence. In A. C. Billings & B. McCullough (Eds.), *The Routledge Handbook of Digital Sport Management* (1st ed., pp. 13-19). Routledge. https://doi.org/10.4324/9781003088899-29

Kushwaha, A. K., & Kar, A. K. (2021). MarkBot–a language model-driven chatbot for interactive marketing in post-modern world. *Information systems frontiers,* 1-18. https://doi.org/10.1007/s10796-021-10184-y

Lee, D., & Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Computers & Education, 191,* 104646. https://doi.org/10.1016/j.compedu.2022.104646

Lesh, R., & Doerr, H. M. (Eds.). (2003). Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In *Beyond constructivism* (pp. 3-33). Routledge.

Lesh R., Hoover M., Hole B., Kelly A., Post T. (2000). Principles for developing thought-revealing activities for students and teachers. In Kelly A., Lesh R. (Eds.), *Handbook of Research in Mathematics and Science Education (pp. 113–149).* Mahwah, NJ: Lawrence Erlbaum and Associates.

Liu, V., & Chilton, L. B. (2022, April). Design guidelines for prompt engineering text-to-image generative models. *In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-23).* https://doi.org/10.1145/3491102.3501825

Lo, C. K., Huang, X., & Cheung, K. L. (2022). Toward a design framework for mathematical modeling activities: An analysis of official exemplars in Hong Kong mathematics education. *Sustainability*, *14*(15), 9757. https:// doi.org/10.3390/su14159757

Lo, L. S. (2023a). The art and science of prompt engineering: A new literacy in the information age. *Internet Reference Services Quarterly,* 1-8. https://doi.org/10.1080/10875301.2023.2227621

Lo, L. S. (2023b). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship, 49(4),* 102720. https://doi.org/10.1016/j.acalib.2023.102720

Lokman, A. S., & Ameedeen, M. A. (2019). Modern chatbot systems: A technical review. *In Proceedings of the Future Technologies Conference (FTC) 2018: Volume 2* (pp. 1012-1023). Springer International Publishing. https://doi.org/10.1007/978-3-030-02683-7_75

Luczak, R., & Erwin, R. (2023). Mathematical modeling: a study of multidisciplinary benefits in the math classroom. *Teaching Mathematics and its Applications: An International Journal of the IMA*, *42*(4), 325-342. https://doi.org/10.1093/teamat/hrac021

Melis, G., Dyer, C., & Blunsom, P. (2017). On the state of the art of evaluation in neural language models. *arXiv preprint arXiv:1707.05589.* https://doi.org/10.48550/arXiv.1707.05589

Nichols, J. A., Herbert Chan, H. W., & Baker, M. A. (2019). Machine learning: applications of artificial intelligence to imaging and diagnosis. *Biophysical reviews, 11,* 111-118. https://doi.org/10.1007/s12551-018-0449-9

Niss, M. (2010). Modeling a crucial aspect of students' mathematical modeling. *In: Lesh, R., Galbraith, P., Haines, C., Hurford, A. (eds) Modeling Students' Mathematical Modeling Competencies.* Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-0561-1_4

OECD. (2013). PISA 2012 *Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy.* OECD Publishing.

OECD. (2018). *PISA 2021 mathematics framework (second draft).*

OpenAI. (2022). Introducing ChatGPT. https://openai.com/blog/chatgpt

Nurhasan, Prahani, B. K., Suprapto, N., & Al Ardha, M. A. (2022). Artificial intelligence research during covid-19 pandemic: contributed to future education. *International Journal of Instruction, 15(3),* 229-248. https://doi.org/10.29333/iji.2022.15313a

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training.*

Reynolds, L., & McDonell, K. (2021, May). Prompt programming for large language models: Beyond the few-shot paradigm. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-7). https://doi.org/10.1145/3411763.3451760

Savelka, J., Agarwal, A., Bogart, C., & Sakr, M. (2023). Large language models (gpt) struggle to answer multiple-choice questions about code. *arXiv preprint arXiv:2303.08033.* https://doi.org/10.48550/arXiv.2303.08033

Sezen Yüksel, N., Sağlam Kaya, Y., Urhan, S. & Şefik, Ö. (2019). *Matematik eğitiminde modelleme etkinlikleri* (Ed: Şenol Dost). Ankara: Pegem Akademi

Short, C. E., & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights, 19,* e00388. https://doi.org/10.1016/j.jbvi.2023.e00388

Spasić, A. J., & Janković, D. S. (2023, June). Using ChatGPT standard prompt engineering techniques in lesson preparation: Role, instructions and seed-word prompts. In 2*023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)* (pp. 47-50). IEEE. https://doi.org/10.1109/ICEST58410.2023.10187269

Susnjak, T. (2022). ChatGPT: The end of online exam integrity?. *arXiv preprint arXiv:2212.09292.* https://doi.org/10.48550/arXiv.2212.09292

Tekin Dede, A., Hıdıroğlu, Ç. N., & Bukova Güzel, E. (2017). Examining of model eliciting activities developed by mathematics student teachers. *Journal on Mathematics Education, 8(2)*, 223–242. https://doi.org/10.22342/jme.8.2.3997.223-242

Urhan, S., & Dost, Ş. (2018). Analysis of ninth grade mathematics course book activities based on model-eliciting principles. *International Journal of Science and Mathematics Education, 16,* 985-1002. https://doi.org/10.1007/s10763-017-9808-4

Vermeersch, A. (2023). Deep learning for K3 fibrations in heterotic/Type IIA string duality. *Nuclear Physics B, 116279.* https://doi.org/10.1016/j.nuclphysb.2023.116279

Wang, S., Scells, H., Koopman, B., & Zuccon, G. (2023). Can ChatGPT write a good boolean query for systematic review literature search?. *arXiv preprint arXiv:2302.03495.* https://doi.org/10.1145/3539618.3591703

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 19(7),* em2286. https://doi.org/10.29333/ejmste/13272

Xia, Q., Chiu, T. K., Zhou, X., Chai, C. S., & Cheng, M. (2022). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence,* 100118. https://doi.org/10.1016/j.caeai.2022.100118

Yong, G., Jeon, K., Gil, D., & Lee, G. (2023). Prompt engineering for zero-shot and few-shot defect detection and classification using a visual-language pretrained model. *Computer-Aided Civil and Infrastructure Engineering, 38(11),* 1536-1554. https://doi.org/10.1111/mice.12954

Zheng, Z., Zhang, O., Borgs, C., Chayes, J. T., & Yaghi, O. M. (2023). ChatGPT Chemistry Assistant for Text Mining and Prediction of MOF Synthesis. *arXiv preprint arXiv:2306.11296.* https://doi.org/10.1021/jacs.3c05819