

Worked-Out Examples with ChatGPT-Based Explanations: A Statistics Activity

Ayşe Tuğba Öner¹

This study aims to prepare statistics activity-targeting comprehension of the arithmetic mean, consisting of worked-out example explanations through ChatGPT for sixth graders; determine situations encountered during the implementation of the activity; ascertain the effect of the activity on students' mathematics problem-solving performance; and obtain students' perception about the activity. The embedded design was used in this study. A one-group pre-test-post-test design was employed to assess the situations experienced during the activity and its effectiveness. Additionally, a generic qualitative inquiry was conducted to understand the students' perceptions of the activity. Thirteen sixth-grade students were involved in the study. In addition to the activity developed, a mathematics performance test and a reflection paper were used to obtain data. As a result, ChatGPT assists students in solving problems through step-by-step explanations. The activity, which utilized ChatGPT-based explanations as a worked-out example, was statistically significant in improving students' mathematics problem-solving performance. Additionally, it was found to be enjoyable, informative, and beneficial for students' learning. However, ChatGPT distorted the given mathematics problem or tables in the problem and displayed incorrect explanations or solved mathematics problems by including a topic that the students had not learned, even though it was stated in the prompt.

Keywords: ChatGPT; statistics activity; worked-out example; cognitive load.

Article Information:

Research Article

Submitted: August 14, 2025

Revised: September 21, 2025

Accepted: October 7, 2025

How to cite:

Öner, A. T. (2025). Worked-Out Examples with ChatGPT-Based Explanations: A Statistics Activity. *Journal of Inquiry Based Activities*, 15(2), 146-169.

¹ İstanbul Medeniyet University, Faculty of Educational Sciences, Department of Mathematics and Science Education, İstanbul, Türkiye. ORCID ID: 0000-0001-9543-6576, E-mail: tugba.oner@medeniyet.edu.tr

Introduction

ChatGPT is an application used by many people in daily life and is a tool that has been frequently used in educational studies and whose effectiveness has been studied (Egara & Mosimege, 2024; Ji et al., 2023; Marcus & Lindvall, 2024; Taani & Alabidi, 2024). ChatGPT is a text-based application that responds to prompts with explanations and, when combined with worked-out example instruction, is suitable for use in mathematics education. Using worked-out examples supports students' learning and increases their problem-solving achievement (Chandler & Sweller, 1991; Cooper & Sweller, 1987; Gökdeniz Tahiroğlu, 2023; Özcan, 2024; Özcan et al., 2018; Sweller & Cooper, 1985). Using worked-out examples is beneficial for reducing students' mental workload by increasing their intrinsic and germane cognitive loads and decreasing their extraneous loads (Ayres, 2012). Using worked-out examples helps students understand topics they struggle with more easily (Özcan, 2024).

Statistics is often perceived as a complex and challenging subject for students (Fitzmaurice et al., 2014; Watts, 1991), especially at higher education level. This may be due to a lack of meaningful learning about statistical concepts at a young age. In addition, while statistical literacy is an important type of literacy that plays a role in daily life (Ben-Zvi & Garfield, 2004; Watson, 2006), research has also shown that students' statistical literacy or reasoning levels are not advanced (Aydın, 2019; Özbay, 2012; Vural, 2020; Yolcu, 2012; Zora, 2019). Therefore, it would be beneficial for students to use worked-out examples that reduce cognitive load while simultaneously facilitating meaningful learning. However, with developing technology, teachers can also utilize generative artificial intelligence (AI) to accelerate this process. As in this study, designing a statistical activity that helps students to comprehend the concept of arithmetic mean, a fundamental concept in statistics, as well as integrating AI and a worked-out example into this activity besides carrying out a process and observing the results presents an important finding in the literature.

Worked-out Examples

Worked-out examples are a type of direct instruction in which a step-by-step solution to a problem or activity is presented (Ayres, 2012; McGinn et al., 2015). The use of worked-out examples in learning differs positively from different types of problem solving, such as solving a large number of examples (Chandler & Sweller, 1991), and is especially recommended as it is effective in learning (Cooper & Sweller, 1987; Sweller & Cooper, 1985; Zhu & Simon, 1987). When worked-out examples are used, students' success in solving similar problems and students' conceptual understanding increase (Booth et al., 2013; Dağseven Emecen, 2020; Gökdeniz Tahiroğlu, 2023; Özcan, 2024; Retnowati et al., 2010; Sweller & Cooper, 1985), minimize errors in problem solving (Özcan et al., 2018; Sweller & Cooper, 1985), lead to successful results in transfer problems (Cooper & Sweller, 1987; Özcan, 2024; Retnowati et al., 2010), and reduce the cognitive load of students (Pass et al., 1994; van Gog et al., 2010, 2011). Paas (1992) compared three different methods (i.e., conventional, worked-out, and worked-out with completion) in his study of measures of central tendency in a statistical subject with high school students and found that the two methods, which used worked examples, were more effective than the conventional method in solving transfer problems and required less mental effort. Using worked-out examples in a statistics course not only improves students' statistical reasoning skills but also increases their confidence in succeeding in learning (Jin et al., 2011). Among the different methods used by university students in an introductory statistics course, worked-out examples were also found to be more effective in learning and to result in lower cognitive load on students (Newman & Decaro, 2019). As a result, the relevance of worked-out examples to the cognitive load theory becomes apparent.

Cognitive Load Theory

Worked-out examples are based on the cognitive load theory (CLT) because the way in which worked-out examples are presented focuses the learner's attention, reduces mental workload, and therefore facilitates learning (Cooper & Sweller, 1987). When students are working on solving a problem, their minds are working on how to solve it, which leads to an increase in their working memory load (Ayres, 2012), but worked-out examples already include the presentation of the solution from each step to the next step at this point, helping students to concentrate on what they really need to learn (Ayres, 2012). Other instructional formats may cause learners to focus on different points rather than focusing on what they need to learn, leading to an increase in cognitive load, which can be experienced in different types of cognitive load. Cognitive load theory deals with the cognitive resources used during learning and problem solving (Chandler & Sweller, 1991) and how instruction should be planned to maximize learners' performance (Sweller, 2012). According to this theory, trying to focus on more than one point at the same time can make learning more difficult because it increases the working memory load (Sweller et al., 2011).

According to CLT, there are three types of cognitive loads: a) intrinsic, b) extraneous, and c) germane. Intrinsic CL is related to the complexity of the information and material to be learned, and in order for it to alter, changes need to be made in the learning itself or in what is learned (Sweller, 2010). It is determined by the quantity of interacting information elements in the task or material being learned (Leppink et al., 2014; Sweller, 2010). Low element interactivity produces low working memory load, when things learned without referring to other elements fall into this group; and someone who is learning a new information, that is, a novice learner, may have to process more than one element at a time, whereas an advanced learner stores information as more than one element integrated and experiences a lower working memory load (Sweller, 2010). In addition to the complexity of the material to be learned, another point that affects working memory load is instructional procedures, and improper organization of these procedures (i.e., nonoptimal) creates extraneous CL (Sweller, 2010). Extraneous CL is undesirable because focusing on an element other than the learning situation results in extraneous CL (Sweller, 2010). Another type of CL is germane CL, which is only related to the learner's characteristics. The importance of this lies in how much the learner prioritizes working memory resources for the learning material (Sweller, 2010). Therefore, when designing learning materials, the aim is to optimize intrinsic CL, reduce extraneous CL, and enable the emergence of germane CL, which will help learners construct their cognitive schema (Leppink et al., 2014). At this point, the importance of worked-out examples emerges because they contribute to schema formation by reducing extraneous CL and allocating space for germane CL (Ayres, 2012).

ChatGPT and Worked-out Examples

ChatGPT is a large language model (LLM), which is a form of artificial intelligence that processes and produces textual data (Jury et al., 2024; Naveed et al., 2025) from OpenAI that has reached a wide audience (Jury et al., 2024; Wright, 2023). In a study where LLMs were used to prepare worked-out examples in computer science education, it was mentioned that worked-out examples were used as an effective teaching method, but these models were used because they were difficult to prepare by instructors (Marcus & Lindvall, 2024). Marcus and Lindvall (2024) concluded in their study that the solved examples produced by LLMs were at a level where some minor modifications were needed and that they produced good-quality examples and results. Similarly, in a study, researchers found that LLMs in an introductory programming course produced good-quality code explanations, especially for novice users; therefore, they developed and used a tool called 'WorkedGen' to produce worked-out examples (Jury et al., 2024). As a result, they concluded that it contained good explanations and was

beneficial for learning (Jury et al., 2024). In a study where a ChatGPT-supported collaborative learning environment was provided, it was found that teacher candidates' STEM learning performance increased and their cognitive load decreased (Ji et al., 2023). ChatGPT is used by mathematics and science teachers to support problem-solving, example generation, and explanations of the subject (Taani & Alabidi, 2024), but it is not a tool that mathematics teachers widely use. However, when it is used, it increases teaching effectiveness, facilitates the understanding of complex concepts, and ensures student engagement (Egara & Mosimege, 2024), and improves students' motivation to learn mathematics (Almarashdi et al., 2024). In a study comparing the answers given by students to questions in chemistry class with the answers given by ChatGPT, it was concluded that ChatGPT gave explanations like worked-out examples in chemistry textbooks and, unlike students, provided detailed, long, and good conceptualization (Clark et al., 2023). Considering these studies, the use of a publicly available LLM platform such as ChatGPT in generating worked-out examples in mathematics, a field where worked-out examples are widely used, can both save teachers' time and be beneficial for students to use for their own learning, not only in the classroom but also in other areas. Likewise, in a study examining how good a mathematics teacher ChatGPT could be, it was determined that all of the lesson plans prepared by ChatGPT included worked-out examples (Cameron & Mesiti, 2024). Since the performance of LLMs depends on the quality of the prompts entered (Hassany et al., 2024), it is important to take into account that the prompts should be entered in accordance with the students' levels if a good performance in solving mathematical problems is desired.

When the importance of worked-out examples on students' achievement, the popularity of ChatGPT usage in education, and its possibility to reduce students' cognitive load during problem-solving activities are considered, preparing an activity including ChatGPT as a tool for worked-out example explanation and determining its effectiveness could give educators information about the future use of this tool in their classroom. Therefore, this study aims to prepare statistics activity-targeting comprehension of arithmetic mean consisting of worked-out example explanations through ChatGPT; determine situations encountered during the implementation of the activity; ascertain the effect of the activity on students' mathematics problem-solving performance; and obtain students' perception about the activity.

Method

The embedded design (Creswell, 2012), a type of mixed methods design, was used in the study. One group pretest-posttest design, a weak experimental research design, was used as the quantitative approach; a generic qualitative inquiry was used to support quantitative results. Generic qualitative inquiry is a study in which known qualitative research approaches (case study, etc.) are not chosen, but instead an answer to the research question is sought using qualitative methods due to their nature (Patton, 2015; Sayın et al., 2022).

Study Group

The study group consisted of 13 sixth-grade students (n=7 male; n=6 female). The study was conducted at a middle school in the Üsküdar district of Istanbul. A convenience sampling technique was used to select participants who were easy to reach (Büyükoztürk et al., 2024; Johnson & Christensen, 2014). The researcher's ease of access to the students, strong communication with the mathematics teacher, and the school administration's emphasis on scientific research practices were all factors considered in selecting this group.

Preparation of the Activity

Worked-out examples are used both to ensure that students learn meaningfully about a topic or concept that they will learn for the first time, and in cases where the teaching process has been carried out before, but students still have difficulty in making sense of it and solving the questions (Özcan, 2024). Before implementing the activity, the mathematics teacher taught a lesson covering the 2018 Mathematics curriculum standard “M.6.4.2.2. (Student) Calculates and interprets the arithmetic mean of a data group” belonging to the data analysis subdomain (Milli Eğitim Bakanlığı [MoNE], 2018, p. 64).

At this point, during the activity preparation process, the mathematics teacher was first contacted to obtain information about the types of questions the students had difficulty with regarding the arithmetic mean, one of the statistics concepts. It was learned that students have more difficulty with far transfer problems than with instructional and near transfer problems that require operational knowledge about the average (Paas, 1992). One of the 6th-grade skill-based questions developed and made available openly by the Ministry of National Education (MoNE) General Directorate of Measurement, Evaluation and Exam Services (MEES) was purposefully selected and presented to the mathematics teacher for his/her opinion. Considering that students struggle to solve such questions, activity questions were prepared by taking inspiration from questions prepared by the Mone MEES and modifying their context.

Another reason for choosing the skill-based questions prepared by the Mone MEES is that the questions in these resources were validated by experts and made available for open access. The questions in the prepared activity sheet were presented to a mathematics teacher and a mathematics education expert for their opinions, and they were asked to evaluate whether the questions were related to the mathematics standard, appropriate for the grade level, and in terms of appearance, thus ensuring content and face validity. After deciding on the suitability of the questions, the instructions for entering the prompts into ChatGPT were prepared in order to obtain worked-out examples with ChatGPT based explanations. There were several considerations at this point. First, the instructions in the activity sheet were prepared assuming that the students had never used ChatGPT before. Thus, instead of entering the question as a text prompt, attention was paid to first taking a picture of the question with the phone camera and then including explanations on how to upload it to ChatGPT. The aim was to prevent students from entering incomplete or incorrect information in the question when using text form, as this could impact the question's solution. Secondly, there was the possibility that ChatGPT would offer a solution using a concept or procedure that was not appropriate for the student's grade level, that is, a concept or procedure that the students had not yet learned (e.g., forming and solving equations).

Before the implementation of the activity, it was known that the subjects of operations with natural numbers and fractions, decimal representation, ratio, algebraic expressions, mass measurement, and data analysis were taught in the classroom. It was notable that 6th-grade students were unfamiliar with setting up and solving first-degree equations with one unknown because this topic is included in the 7th-grade mathematics curriculum in Türkiye. To prevent the possibility of ChatGPT providing a solution that is not appropriate for the student's grade level, special attention was paid to informing the students that they were 6th-grade students and did not know how to form equations in the ChatGPT prompt.

The third issue was that the prompt entry into LLMs should be arranged in a way that serves the purpose (Hassany et al., 2024). At this point, due to the need for the explanations on the ChatGPT screen to be explanations like the worked-out examples, in the prompt entered into ChatGPT, it was necessary to be careful to organize the entered prompt in case ChatGPT made the explanation of the

problem short, incomplete, or solved the question directly. Therefore, another point that was taken into consideration in the instructions given to ChatGPT was the request to explain the problem verbally, step by step, to ensure that it included explanations of the solution. Finally, an instruction was given on the activity sheet asking the students to examine the explanations and solution steps made by ChatGPT in detail. Before the activity was implemented in the classroom, the students' mathematics teachers were contacted and informed that the students would use their mobile phones in mathematics class. In addition, since students needed to create an account to upload the photo of the mathematics problem to ChatGPT, and because students might not have used an account before, the researcher prepared email addresses and passwords the day before.

Ethics committee approval was obtained for the study with the decision of the Social and Human Sciences Ethics Committee of Istanbul Medeniyet University dated 07/01/2025 and numbered 2025/06.

Implementation of the Activity

Before the activity was implemented, a pre-test consisting of an open-ended question inspired by a skill-based question prepared by MoNe MEES was administered in order to determine the effect of the application on the students' performance in the far-transfer questions. After the pre-test application, students were asked what they knew about ChatGPT and whether they had used it before, as an introductory phase. Afterwards, students were given their phones and instructed to follow only the activity sheet's instructions, using their phones only for ChatGPT. Some students had ChatGPT installed on their mobile phones and used the application for transactions, while others accessed the ChatGPT screen through a web browser. The version of ChatGPT the students used was the free version of ChatGPT-4o. All students logged in with their email addresses and were ready to follow the instructions on the activity sheet. Activity sheets (Figures 1-2) were distributed, and the researcher first introduced the steps and prompts on the activity sheet (see Appendix A for Turkish). The students were then left alone with their mobile phones to practice the steps on the activity sheet. After the first activity was completed, a second activity sheet was distributed, and students repeated the same process.

Students experiencing technical difficulties were addressed individually and resolved. After the implementation of the activity, a post-test was administered to determine students' performance on far-transfer questions related to the concept of arithmetic mean using worked-out examples with ChatGPT-based explanations. Finally, the in-class implementation was concluded by distributing reflection papers to determine students' perceptions of the activity.

CHATGPT SUPPORTED STATISTICS ACTIVITY

Mehmet, who went to the market with his mother, asked her how many kilograms of walnuts they would need for the dessert they were going to prepare. His mother replied: “From 1000 grams of walnuts in shells, 500 grams of walnut kernels are obtained.” Based on her recollection, Mehmet’s mother noted down the amounts of shelled walnuts and the corresponding amounts of walnut kernels obtained she had bought during three different shopping trips, and organized them into the following table.

Shopping trips	Shelled walnuts (gr)	Walnut kernels (gr)
1 st -trip	800	
2 nd -trip		300
3 rd -trip		

Mehmet calculated from his mother’s grocery receipts that the arithmetic mean of the amounts of walnuts in their shells she bought across three different shopping trips was 600 grams. Accordingly, how many grams of walnut kernels did they obtain from the shelled walnuts purchased on the 3rd trip?

Instructions: To get help from ChatGPT in solving this problem, follow the steps presented to ChatGPT below in the given order.

1) First, open a session in ChatGPT. Type the following prompt into ChatGPT: “Hello, I am a sixth-grade student and I do not know how to set up equations. I will ask you a question, but I do not want you to solve it. I only want you to explain verbally, step by step, how the problem should be solved.” Now press the Enter key or tap the up-arrow icon  to send this message.

2) Now you need to upload the question to ChatGPT. To do this:

- First, tap the icon  located at the lower-left corner of ChatGPT’s menu.
- Then use your phone’s camera to take a photo of the question.
- Wait for the photo to finish uploading, and when the upload is complete press Enter or tap the up-arrow icon .

3) You will now see ChatGPT’s verbal explanation of how the problem should be solved. Read the explanation carefully.

4) To see the numerical solution, type: “Now I want you to show the solution to the problem along with the explanations, but please do not include any equations.” Then press Enter or tap the up-arrow icon  to send this message.

5) You will see on the screen ChatGPT’s explanation of how the problem is solved together with the solution steps. Carefully review the explanations and the steps of the solution.

Figure 1

ChatGPT-Based Statistics Activity Sheet I

CHATGPT SUPPORTED STATISTICS ACTIVITY

Mine decided to bake four different types of cakes for the friends she invited to her home. In the cakes she bakes, the ratio of flour to oil she uses does not change, regardless of the type of cake. The measurement she determined for each cake type is as follows: If 1 cup of oil is used, then 4 cups of flour must be used. Mine attempted to organize the amounts of flour and oil she used for the different types of cakes into the following table.

Cake types	Flour (cup)	Oil (cup)
Chocolate		2
Carrot		1/2
Raisin	2	
Plain		

Mine calculated that the arithmetic mean of the oil amounts used in the four cakes was 1 cup. Accordingly, how many cups of flour did she use for the plain cake?

Instructions: To get help from ChatGPT in solving this problem, follow the steps presented to ChatGPT below in the given order.

1) First, open a session in ChatGPT. Type the following prompt into ChatGPT: “Hello, I am a sixth-grade student and I do not know how to set up equations. I will ask you a question, but I do not want you to solve it. I only want you to explain verbally, step by step, how the problem

should be solved.” Now press the Enter key or tap the up-arrow icon  to send this message.

2) Now you need to upload the question to ChatGPT. To do this:

- First, tap the icon  located at the lower-left corner of ChatGPT’s menu.
- Then use your phone’s camera to take a photo of the question.
- Wait for the photo to finish uploading, and when the upload is complete press Enter or

tap the up-arrow icon .

3) You will now see ChatGPT’s verbal explanation of how the problem should be solved. Read the explanation carefully.

4) To see the numerical solution, type: “Now I want you to show the solution to the problem along with the explanations, but please do not include any equations.” Then press Enter or tap the up-arrow icon  to send this message.

5) You will see on the screen ChatGPT’s explanation of how the problem is solved together with the solution steps. Carefully review the explanations and the steps of the solution.

Figure 2

ChatGPT Based Statistics Activity Sheet II

Data Collection Tools

Mathematics Performance Test

This test, consisting of an open-ended question, was created by the researcher to measure the effect of the worked-out examples with ChatGPT-based explanations activity on students' mathematics problem-solving performance about the arithmetic mean concept. While creating this test, it was taken into account that students had difficulty with far-transfer questions regarding the concept of the arithmetic mean. The question regarding the acquisition of calculating and interpreting the arithmetic mean in the skill-based tests for 6th graders, prepared as support material by MoNE MEES, was taken into consideration, and an open-ended question was created inspired by the question in this multiple-choice test. In the 2023 education vision, the programs aimed to develop students' high-level skills and to provide multiple materials that will enable the development of these skills (MoNE, 2019). For this purpose, skill-based questions, which start with a scenario or situation and are supported by pictures, tables, or examples (Selçuk, 2020), and which require the analysis of which mathematical knowledge and skills are used in the situation presented, in addition to the procedural skills, have become widespread since 2019 (Gürbüz Us & Ercan Güven, 2022). The reason for using skill-based questions, which are openly available to MoNE MEES, is that experts prepared the questions, and they have been validly and reliably prepared. In order to obtain the content and face validity of the question, whose context was preserved but whose data was changed and turned into an open-ended question by the researcher, a table of specifications was created, including a single standard (i.e., learning outcome), and opinions were received from a middle school mathematics teacher and an associate professor of mathematics education. When calculating the students' test scores, a holistic scoring rubric was used as follows: no solution or incorrect solution, 0 points; partial solution using the given information in the problem but without using the arithmetic mean, 5 points; complete and correct solution using the given information in the problem and the arithmetic mean, 10 points.

Reflection Paper

A four-question reflection paper was prepared to gather students' perceptions on statistics activities consisting of worked-out examples with ChatGPT-based explanations. The reflection paper, collected from students at the end of the activity, aimed to identify students' perceptions of the activity's effectiveness, the application's practicality, and any challenges during the implementation.

Data Analysis

In order to examine the effects of statistics activities consisting of worked-out examples with ChatGPT-based explanations on students' mathematics problem-solving performance, the distributions of pre-test and post-test scores were examined regarding normality. According to the Shapiro-Wilk test for normality, it was found that both pre- and post-test scores were not normally distributed, $p=.001$ and $p=.009$, respectively. Therefore, to analyze the differences in students' pre- and post-test scores, the Wilcoxon Signed Rank test was used as a non-parametric test for statistical hypothesis testing. Additionally, the responses on the reflection papers collected from students at the end of the activity were analyzed through content analysis. The researcher and another mathematics education researcher initially created codes using open coding, and a 95% agreement rate was found. Two researchers then identified categories and finalized the themes.

Results

In this section, the assistance and shortcomings of statistics activity consisting of worked-out examples with ChatGPT-based explanations, the effect of implementation of this activity on students'

mathematics problem-solving performance, and the students' perceptions of this process are presented under three headings.

Assistance and Shortcomings of ChatGPT-Based Explanations

The most important point that attracted attention during the activity was to ensure that the photos were clear and readable by the application when entering them into ChatGPT. Most of the students did not have any problems with the explanations given by ChatGPT regarding the solution of the mathematics problem if they first took a photo of the mathematics problem with their phones, edited the photo according to the instructions on the activity sheet, that is, excluding the parts other than the mathematics problem if any, and uploaded it to the application. In this case, most of the students examined worked-out examples with ChatGPT-based explanations for both activities without any problems and took notes on their activity sheets. A sample screenshot of the solution for the first activity, taken from a student's ChatGPT account, is presented (see Appendix A). In addition, it included explanations that will recall the student's schema regarding the mathematical concept in the mathematics problem with a simpler, similar question (see Figure 3). These were the findings about the assistance of ChatGPT-based explanations.

Bu çözüm mantığını daha iyi anlayabilmen için sana yardımcı olacak kısa bir soru:
Bir şeyin yarısı 600 gram ise, tamamı kaç gramdır?
 İşte bu tür düşüncelerle, bu soruları çok rahat çözebilirsin!

Note. To help you better understand this reasoning, here's a short guiding question: *If half of something is 600 grams, how many grams is the whole?* With this kind of thinking, you can solve these questions very easily!

Figure 3

Example of Simplification of the Problem by ChatGPT

However, during the activity, some students reported difficulties with the solution explanations provided when they uploaded mathematics problems to ChatGPT. These difficulties were: a) issues that occurred when the photo uploaded to the application for this activity is not uploaded according to the instructions stated in the activity sheet; b) ChatGPT did not read the table in the mathematical problem correctly, even though the photo was uploaded according to the instructions; c) it produced some explanations that did not exist in the mathematical problem; and d) although the instruction specifically asks not to establish equations in solving mathematical problems, it contained explanations for forming and solving equations. Figure 4 shows a screenshot of the ChatGPT screen of a student who uploaded an image to ChatGPT without following the instructions. Because this student did not upload an image according to the instructions, ChatGPT rendered the table in the mathematics problem incorrectly.

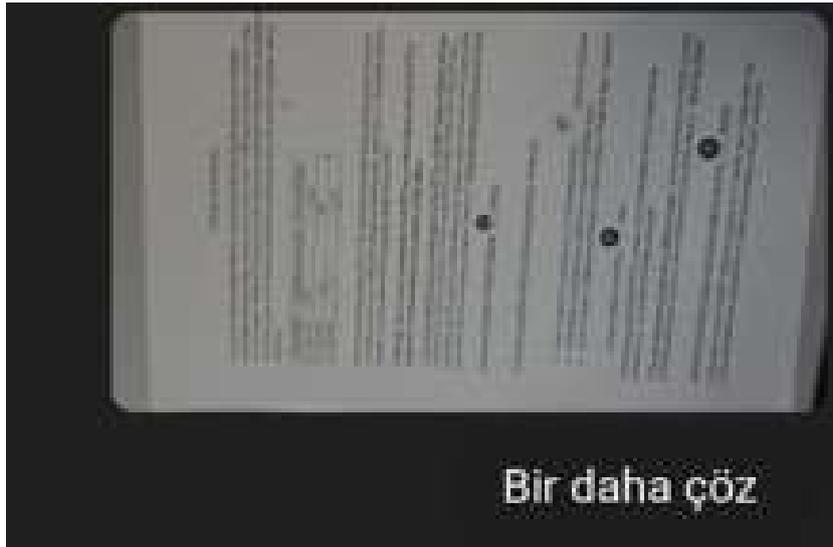


Figure 4

A Screenshot of ChatGPT with the Wrong Photo Uploaded

Even though students uploaded the correct photo of the mathematics problem, ChatGPT failed to process the problem correctly. For example, two screenshots showing two different students' incorrectly tabulated mathematics problems are presented in Figure 5.

Soru Özeti:
 Mehmet 3 alışveriş yapmış. Her alışverişte kabuklu ceviz almış ve bunlardan çıkan ceviz içi miktarlarını not etmiş.

Verilen tablo:

Alışveriş	Kabuklu ceviz (gr)	Ceviz içi (gr)
1. alışveriş	800	300
2. alışveriş	900	360
3. alışveriş	600	?

Aşağıdaki tablo verilmiş:

Alışveriş	Kabuklu ceviz (gr)	Ceviz içi (gr)
1. alışveriş	800	300
2. alışveriş	1000	500
3. alışveriş	?	600

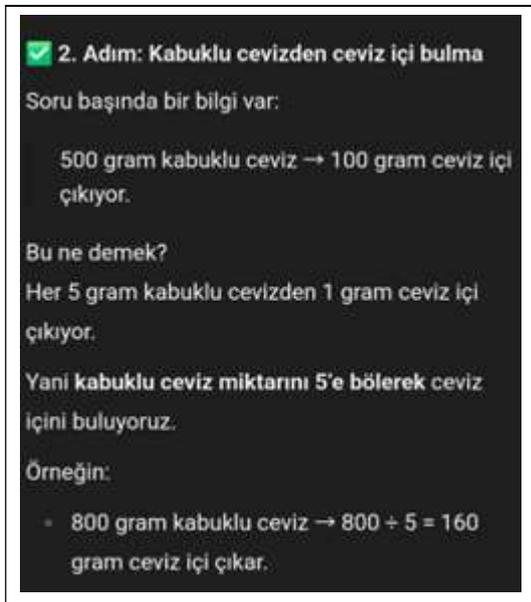
• Bu cevizleri kırıp ne kadar ceviz içi çıktığını ölçmüş.

Note. Translation of the first figure: Mehmet went shopping three times. Each time, he bought walnuts in their shells and noted down the amount of walnut kernels he obtained. The given table: 1st Shopping- 800 gr walnuts in their shells, 300 gr walnut kernels; 2nd shopping- 900 gr walnuts in their shells, 360 gr walnut kernels; 3rd shopping- 600 gr walnuts in their shells, ? gr walnut kernels. Translation of the second figure: He cracked these walnuts and measured how many grams of walnut kernels came out. The following table is given: 1st Shopping- 800 gr walnuts in their shells, 300 gr walnut kernels; 2nd shopping- 1000 gr walnuts in their shells, 500 gr walnut kernels; 3rd shopping- ? gr walnuts in their shells, 600 gr walnut kernels.

Figure 5

Examples of the Table in the Mathematics Problem Rendered by ChatGPT

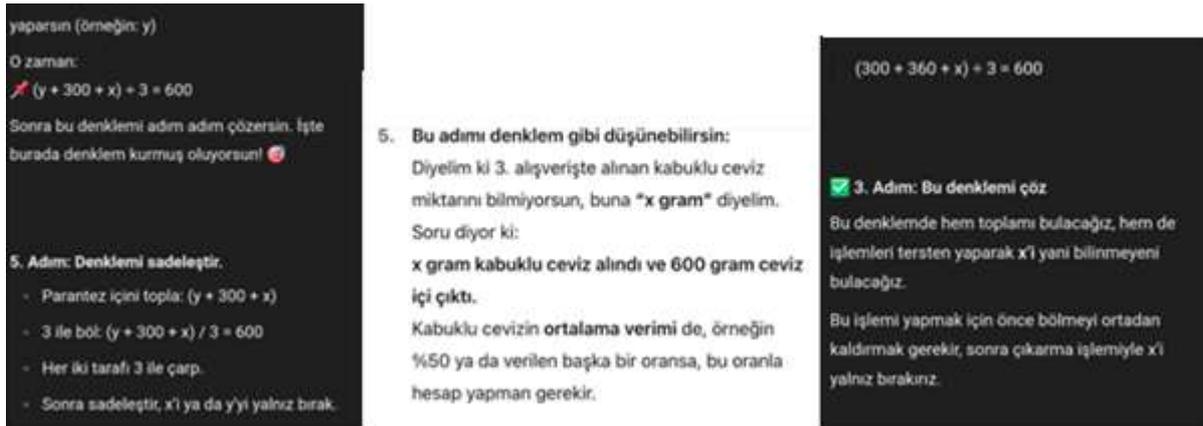
The table in the mathematical problem in Figure 1 indicates that the first purchase includes 800 grams of walnuts in shell, and the second purchase includes 300 grams of walnut kernel. However, as seen in Figure 5, ChatGPT wrote 300 grams of walnut kernel in the first purchase row in both examples. Another numerical value in the problem, 600, which is the arithmetic mean of in-shell walnuts obtained from three purchases, was randomly placed in the third purchase row in both examples, as seen in Figure 5. It appears that the other numbers in the problem (i.e., 1000 and 500) are incorrectly placed on the line representing the second purchase. It was noticed that the numbers 900 and 360 in the other student's solution were processed randomly. Another shortcoming was that the mathematics problem was detected incorrectly by ChatGPT (see Figure 6). Although the explanations proceed step by step, just like in the worked-out examples, and the meaning of each step is defined, as seen in Figure 6, ChatGPT's misunderstanding of the mathematical problem caused the student to examine an incorrect solution. The most significant shortcoming was that, despite being asked to present the solution without establishing an equation, as specified in the activity sheet and indicated as a prompt, ChatGPT actually set up and solved an equation (see Figure 7).



Note. Step 2: Finding walnut kernels from walnuts in their shells. At the beginning of the problem, we are given some information: 500 grams of walnuts in shells → 100 grams of walnut kernels. What does this mean? It means that for every 5 grams of walnuts in shells, you get 1 gram of walnut kernels. In other words, we find the kernel amount by dividing the shelled walnut amount by 5. For example: 800 grams of walnuts in shells → $800 \div 5 = 160$ grams of walnut kernels.

Figure 6

A Screenshot of the Mathematics Problem Being Incorrectly Detected by ChatGPT



Note. Translation of the first figure: Then you solve this equation step by step. This is the point where you are actually forming an equation! Step 5: Simplify the equation. Add what's inside the parentheses: $(y + 300 + x)$, divide by 3: $(y + 300 + x) / 3 = 600$, then multiply both sides by 3, and finally simplify to isolate either x or y . Translation of the second figure: You can think of this step as working with an equation: Let's say you don't know the amount of walnuts in shells bought in the 3rd shopping trip. Let's call it x grams. The problem states that x grams of walnuts in shells were bought, and 600 grams of walnut kernels were obtained. Since the average yield of walnuts (e.g., 50% or another given ratio) is provided, you need to calculate using that ratio. Translation of the third figure: Step 3: Solve this equation. In this equation, we will both find the total and use inverse operations to solve for x , the unknown. To do this, we first need to eliminate the division, and then, by subtraction, isolate x .

Figure 7

Three Different ChatGPT Screenshots Including Equations in the Solution

During the implementation of the activity, any of these results reported as shortcomings of ChatGPT were identified in two situations. The first instance occurred when students themselves identified inaccuracies in the answers provided by ChatGPT. The second instance involved the researcher and mathematics teacher noticing discrepancies while examining the students' phone screens or the activity sheets on which students had written down the answers received from ChatGPT. When encountering the different types of shortcomings mentioned above, the researcher and mathematics teacher intervened and had the students write new prompts to address the issues they encountered. The researcher provided these prompts to the students, and the process of writing new prompts continued until the desired and correct answer was obtained, with each student's ChatGPT response being checked individually. The researcher dictated new prompts to the students after encountering issues with ChatGPT because the students were inexperienced in entering the necessary prompts to achieve the desired results. Another consideration was the cognitive load the students would expend to understand whether the answer to the mathematics problem on the screen was correct after each prompt entered. Since the aim of this study was not to intentionally increase students' extraneous cognitive load, the researcher provided prompts so that the desired solution of the mathematical problem would appear, in order to avoid the possibility that each incorrect answer provided by ChatGPT could potentially increase the student's extraneous cognitive load.

The Effect of Statistics Activities

To determine the effect of the statistics activity on students' mathematics problem-solving performance, a non-parametric test was conducted due to non-normality in both test scores. The

number of students with incorrect/unsolved, partial, and complete solutions in the pre- and post-tests is presented in Table 1. Descriptive statistics for students' pre- and post-test scores are shown in Table 2.

Table 1

Number of Students According to Holistic Rubric Before and After Statistics Activities

	Incorrect/No solution	Partial	Complete
Before activity	8	3	2
After activity	2	7	4

Table 2

Descriptive Statistics of Students' Pretest-Posttest Scores

	Mean	SD	Skewness	Kurtosis
Pre-test	2.69	3.88	1.11	- .15
Post-test	5.76	3.44	- .20	- .49

The Wilcoxon signed-rank test was used to compare students' pre- and post-test results on the far-transfer question about the arithmetic mean concept. According to the results of this test, there was a statistically significant increase ($p=0.046$) in students' post-test scores compared to their pre-test scores (see Table 3).

Table 3

Wilcoxon Signed Rank Test Results

Posttest-Pretest	Ranks	N	Mean Rank	Sum of Ranks	Z	p
	Negative Ranks	1	2.5	2.5	-1.994	0.046
	Positive Ranks	6	4.25	25.5		
	Ties	6				
	Total	13				

Students' Perception of Activities

Students' perceptions about the activity were grouped under three themes: benefits on learning, attributes of application (i.e., ChatGPT), and positive emotions towards the activity.

Benefits on Learning

Students responded differently regarding the activity's benefits for learning. When examining the learning benefits, the first thing that stands out was students' statements regarding the characteristics peculiar to the worked-out example. In the activity, students found it useful that ChatGPT presented the solution of the mathematical problem step by step, as in the worked-out examples. They stated that this made the solution to the mathematical problem easier to understand. Another advantage of worked-out examples was that they help students identify the steps they do not understand in similar questions they have previously solved. For example, two different students' opinions on these issues

are as follows: *"I would like to understand the questions step by step rather than not understanding and leaving them blank. I would like to see what I did not understand"* and *"...if the questions were solved this way (in other courses), we could understand the subject more easily"*.

Another aspect of ChatGPT's beneficial impact on learning was its use as a scaffolding tool. The fluidity of ChatGPT's narrative, its proficiency in summarizing answers when necessary, its ability to explain challenging topics, and its capacity to facilitate tasks that students find difficult all demonstrated that ChatGPT is a tool with distinct features that support students' learning. Students' answers such as *"...made it easier for me to answer the question I couldn't answer"*, *"the explanation and solutions of the questions are very fluent and very good"* and *"...explains in detail when we have difficulty"* showed that the application scaffolds learning.

Attributes of Application

Since the use of the ChatGPT application was the major issue in this activity, the advantages and disadvantages that the students experienced in using the application during the activity formed two categories of this theme. Students listed ChatGPT's advantages as providing practical solutions to mathematical problems, making calculations easier for students, providing informative explanations about the topic or question, and fulfilling the entered prompts as they are. One student's statement, *"...it could be very useful in daily life"* suggested that students can use this tool for different purposes. However, students also pointed out the problems encountered during the implementation of the activity. Student's statement, *"I think some corrections should be made, because it didn't understand the last question no matter what we did"* caused the students to focus on the problems arising from ChatGPT during the activity and stray away from the purpose of the activity, as seen in the shortcomings revealed in the first finding of this study. The students' concerns included ChatGPT's incorrect perception of mathematical problems and its mistakes in solutions. In addition, the fact that the free version used during the activity only allowed a limited number of prompts caused the students to not complete the learning process, which was expressed by the students as a disadvantage of the application.

Positive Emotions Towards Activity

Instead of a traditional mathematics lesson, students engaged in a statistics activity using ChatGPT, a popular mobile phone application, which fostered positive feelings about the activity. Students expressed their pleasure for the activity with phrases like, *"I thought it was a good and educational activity."* Furthermore, many students felt enjoyment during the activity, as evidenced by statements like, *"I thought it was very fun and enjoyable"* and *"the lesson was enjoyable"*.

Conclusion and Discussion

The first aim of this study was to prepare statistics activities targeting comprehension of the arithmetic mean, consisting of worked-out examples with ChatGPT-based explanations. The main reason for creating such an activity was to reduce students' cognitive load. Especially when students are exposed to worked-out examples as instructional material, it alleviates their extraneous cognitive load and increases intrinsic and germane cognitive load (Leppink et al., 2013). However, preparing such activities for each topic could be exhausting for teachers, so they can get assistance from artificial intelligence (AI). Since ChatGPT is a very practical and popular AI tool, creating activity sheets with its assistance would be timesaving for educators. Thus, combining an efficient instructional method in mathematics, which is worked-out examples in this study, with AI would be supportive for both teachers in terms of reducing their workload, and students' in terms of increasing their IL and GL.

During the implementation of the activities, both satisfactory and inadequate results were obtained. When ChatGPT followed given prompts properly, which were written as instructions on the activity sheet, and provided correct solutions to mathematics problems with step-by-step explanations, it was a powerful tool to use in the classroom. Students found ChatGPT's explanations to be informative and helpful in understanding the concept. Certainly, one of the things that students thought was beneficial for their learning was the structure of explanations, which were worked-out examples of a type. According to students' perceptions, another finding was that ChatGPT plays the role of a scaffolding tool for learning. Instead of teacher assistance, ChatGPT's explanations resemble just-in-time adult assistance (Gonulal & Loewen, 2018), thereby providing scaffolding that improves students' mathematical problem-solving skills (Zhou et al., 2025). This finding suggests that educators can utilize generative AI to support teaching. On the other hand, it was also found that ChatGPT had some shortcomings. For instance, ChatGPT produced explanations that did not exist in the mathematical problem and set up equations to solve mathematical problems, even though it was instructed not to include an equation in the prompt. The latter one is a problematic issue because the statistics activity was designed for 6th graders, whose mathematics curriculum does not cover equations (MoNE, 2018). Exposing students to an answer that is not the purpose of the activity and that includes a topic they do not know in the solution stages of a mathematical problem will increase the students' cognitive load and especially cause an increase in EL. However, the purpose of this activity was to reduce students' EL and CL. This shortcoming of ChatGPT was counterproductive to the purpose of the activity. To prevent students from encountering this situation in the classroom environment, teachers can use ChatGPT to prepare worked-out examples before instruction as a preparation step, and after checking its answers, they can use its explanations as a worked-out example task sheet. This will keep ChatGPT's answer under control and will also help alleviate the teacher's workload.

Although some students encountered the above-mentioned issues during the activity's implementation, others did not experience any problems. This suggests that the activity was beneficial to some students. The other aim of this study was to determine the effectiveness of the activity, which was statistically efficient in this specific sample. Its effectiveness on students' mathematical problem-solving performance was an important finding, which could mean that the activity also enhance students IL. If students did not encounter any difficulties or misunderstandings with ChatGPT's interpretation, they found it informative and helpful for understanding the mathematical topic. Even though they realized that ChatGPT has some shortcomings, they still enjoyed the activity. Enjoyment is an important emotion that has a role in students' achievement (Pekrun et al., 2002; Schukajlow et al., 2012). Thus, creating enjoyable activities for instructional purposes, as in this study, especially involving applications or games, would be effective for learning and motivate them towards the course.

One of the aims of this study was to understand the students' thoughts on the activity process. Although most of them found this process pleasant and enjoyable, they also complained about the disadvantages of ChatGPT. One of the significant disadvantages was the deterioration of the mathematics problem by ChatGPT. Students had to identify a table that was incorrect in the explanations provided by ChatGPT. Another problem was the limited access of students to GhatGPT-4o with the free version. When they encountered a problem in the explanation and wanted to guide it through more prompts, they were unable to do so after a few trials. Although this is a limitation, the researcher made sure that the process was completed by using additional accounts so that the students would not be affected during the implementation.

Sample size is one of the limitations of this study. Because one of the aims of this study is to prepare the activity sheet and determine the positive and negative aspects encountered during the activity, a

small sample size was initially considered sufficient to achieve this goal. However, since the study had a weak experimental design without a control group and was conducted on a small sample group, which poses a threat to external validity, it should not be forgotten that the results have limited generalizability in a quantitative aspect when interpreting them. In studies where activity sheets similar to the one developed in this study are developed or where worked-out examples with ChatGPT-based explanations are used as an instruction method, it may be recommended to work with experimental designs having control groups with larger sample sizes. Another limitation is that the study focused only on the concept of the arithmetic mean. The reason for this is the expectation that the sixth grade students should have learned this concept at previous grade levels and were capable of answering far-transfer questions. Future research could further investigate the impact of teaching worked-out examples with ChatGPT-based explanations on students' statistical skills at different grade levels, with activities designed for other statistical concepts and both near and far transfer questions. Using free version of ChatGPT, which the students also mentioned this issue in their perceptions, was a limitation as well. The use of the paid version could be investigated to determine whether the problems encountered in this study (i.e., producing explanations that did not exist in the mathematical problem and setting up equations to solve mathematical problems) would be encountered and whether there would be differences in students' opinions about the application when they felt more comfortable with the number of prompts.

In conclusion, using AI tools, specifically ChatGPT, to enhance students' mathematics learning and performance brings with it both advantages and limitations. Due to students' enjoyment and pleasantness during this activity, using ChatGPT's motivational factor may counterbalance the limitations and compensate for its problematic results. However, this is a matter requiring extreme caution, like walking on thin ice, and educators are advised to carefully prepare their activities in this regard.

Conflict of Interest

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

References

- Almarashdi, H. S., Jarrah, A. M., Khurma, O. A., & Gningue, S. M. (2024). Unveiling the potential: A systematic review of ChatGPT in transforming mathematics teaching and learning. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(12), 1-14. <https://doi.org/10.29333/ejmste/15739>
- Aydın, Ş. (2019). *Sekizinci sınıf öğrencilerinin merkezi eğilim ölçülerine yönelik istatistiksel düşüncelerinin incelenmesi* (Publication No. 644516) [Master's thesis, Mersin Üniversitesi]. Ulusal Tez Merkezi.
- Ayres, P. (2012). Worked example effect. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 3467–3471). Springer. <https://doi.org/10.1007/978-1-4419-1428-6>
- Ben-Zvi, D., & Garfield, J. (Eds.). (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Springer Netherlands. <https://doi.org/10.1007/1-4020-2278-6>
- Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K. J. (2013). Using example problems to improve student learning in algebra : Differentiating between correct and incorrect examples. *Learning and Instruction*, 25, 24–34. <https://doi.org/10.1016/j.learninstruc.2012.11.002>
- Büyükoztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2024). *Eğitimde bilimsel araştırma yöntemleri* (36th ed.). Pegem Akademi.

- Cameron, S., & Mesiti, C. (2024). What kind of mathematics teacher is ChatGPT? Identifying the pedagogical practices preferred by generative AI tools when preparing lesson plans. In J. Višňovská, E. Ross, & S. Getenet (Eds.), *Surfing the waves of mathematics education. Proceedings of the 46th annual conference of the Mathematics Education Research Group of Australasia* (pp. 135–142). Gold Coast: MERGA. <https://files.eric.ed.gov/fulltext/ED661134.pdf>
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332.
- Clark, T. M., Anderson, E., Dickson-Karn, N. M., Soltanirad, C., & Tafini, N. (2023). Comparing the performance of college chemistry students with ChatGPT for calculations involving acids and bases. *Journal of Chemical Education*, 100(10), 3934–3944. <https://doi.org/10.1021/acs.jchemed.3c00500>
- Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, 79(4), 347–362.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.
- Dağseven Emecen, D. (2020). Öğrenme güçlüğü gösteren öğrencilere problem uygulamasının etkililiği. *Karadeniz Uluslararası Bilimsel Dergi*, 48, 81–101. <https://doi.org/10.17498/kdeniz.827222>
- Egara, F. O., & Mosimege, M. (2024). Exploring the integration of artificial intelligence-based ChatGPT into mathematics instruction: Perceptions, challenges, and implications for educators. *Education Sciences*, 14(7), 1-20. <https://doi.org/10.3390/educsci14070742>
- Fitzmaurice, O., Leavy, A., & Hannigan, A. (2014). Why is statistics perceived as difficult and can practice during training change perceptions? Insights from a prospective mathematics teacher. *Teaching Mathematics and Its Applications*, 33(4), 230–248. <https://doi.org/10.1093/teamat/hru010>
- Gökdeniz Tahiroğlu, N. (2023). *Kesirler konusundaki farklı çözüm yollarını içeren çözümlü örneklerin bir web 2.0 aracı ile açıklanmasının beşinci sınıf öğrencilerinin başarısına, matematik tutumuna ve motivasyonlarına etkisinin incelenmesi*. (Publication No. 791038) [Master's thesis, İstanbul Medeniyet Üniversitesi]. Ulusal Tez Merkezi.
- Gonulal, T., & Loewen, S. (2018). Scaffolding technique. In J. I. Lontas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1–5). Wiley. <https://doi.org/10.1002/9781118784235.eelt0180>
- Gürbuz Us, R., & Ercan Güven, A. N. (2022). Millî eğitim bakanlığı tarafından hazırlanan beceri temelli türkçe sorularının yenilenmiş Bloom taksonomisi açısından ele alınması. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, 35(2), 426–445. <https://doi.org/10.19171/uefad.1078375>
- Hassany, M., Brusilovsky, P., Ke, J., Akhuseyinoglu, K., Balajice, A., & Narayanan, L. (2024). Human-AI co-creation of worked examples for programming classes. *Joint Proceedings of the ACM IUI Workshops*, 1–12.
- Ji, Y., Zou, X., Li, T., & Zhan, Z. (2023). The effectiveness of ChatGPT on pre-service teachers' STEM teaching literacy, learning performance, and cognitive load in a teacher training course. *ACM International Conference Proceeding Series*, March, 16–22. <https://doi.org/10.1145/3637907.3637948>
- Jin, L., Kim, Y. J., McGhee, M., & Reiser, R. (2011). Statistical reasoning skills and attitude: the

- effect of worked examples. In M. Simonson (Ed.), *34th annual proceedings of the annual convention of the association for educational communications and technology* (Vol. 1, pp. 105–110).
- Johnson, R. B., & Christensen, L. (2014). *Educational research: Quantitative, qualitative, and mixed approaches* (5th ed.). Sage.
- Jury, B., Lorusso, A., Leinonen, J., Denny, P., & Luxton-Reilly, A. (2024). Evaluating LLM-generated worked examples in an introductory programming course. *Australian Computing Education Conference (ACE 2024)*. <https://doi.org/10.1145/3636243.3636252>
- Leppink, J., Paas, F., Gog, T. Van, Vleuten, C. P. M. Van Der, & Merriënboer, J. J. G. Van. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, 30, 32–42. <https://doi.org/10.1016/j.learninstruc.2013.12.001>
- Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(4), 1058–1072. <https://doi.org/10.3758/s13428-013-0334-1>
- Marcus, S., & Lindvall, V. (2024). *Using large language models to generate worked examples of cs2-level programming questions*. KTH Royal Institute of Technology.
- McGinn, K. M., Lange, K. E., & Booth, J. L. (2015). A worked example for creating worked examples. *Mathematics Teaching in the Middle School*, 21(1), 26–33.
- Milli Eğitim Bakanlığı [MoNE]. (2018). *Matematik dersi öğretim programı [Mathematics curriculum]*.
- MoNE. (2019). *2023 vizyonu*. <https://tegm.meb.gov.tr/www/2023-vizyonu/icerik/23>
- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2025). A comprehensive overview of large language models. *ACM Transactions on Intelligent Systems and Technology*, 16(5)1-72. <https://doi.org/10.1145/3744746>
- Newman, P. M., & Decaro, M. S. (2019). Learning by exploring : How much guidance is optimal ? *Learning and Instruction*, 62, 49–63. <https://doi.org/10.1016/j.learninstruc.2019.05.005>
- Özbay, S. (2012). *İnformel çıkarsamalı akıl yürütmede öğrencilerin örneklem hakkındaki akıl yürütme ve düşünme süreçleri* (Publication No. 325765) [Master's thesis, Pamukkale Üniversitesi]. Ulusal Tez Merkezi.
- Özcan, Z. Ç. (2024). Evaluating the impact of the WEI4S instructional approach on middle school students' algebraic problem-solving skills. *Education Sciences*, 14(102), 1–20. <https://doi.org/https://doi.org/10.3390/educsci14010102>
- Özcan, Z. Ç., Kılıç, Ç., & Obalar, S. (2018). Öğrencilerin matematikteki hatalarını belirleme ve gidermede açıklayıcı ipuçlarıyla desteklenmiş çözümlü örnekler. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 45, 1–22. <https://doi.org/10.21764/maeuefd.322223>
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429–434. <https://doi.org/10.1037/0022-0663.84.4.429>
- Pass, F. G. W. C., Van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79, 419–430.
- Patton, M. Q. (2015). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). Sage.

- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Positive emotions in education. In E. Frydenberg (Ed.), *Beyond coping: Meeting goals, visions, and challenges* (pp. 149–174). Elsevier.
- Retnowati, E., Ayres, P., & Sweller, J. (2010). Worked example effects in individual and group work settings. *Educational Psychology, 30*(3), 349–367. <https://doi.org/10.1080/01443411003659960>
- Sayın, S., Özdemir, Ü., & Öner, A. T. (2022). Mathematics teachers' instructional explanations and estimation skills about area measurement. *Kırşehir Eğitim Fakültesi Dergisi, 23*(Special Issue), 84–127. <https://doi.org/10.29299/kefad.997703>
- Schukajlow, S., Leiss, D., Pekrun, R., Blum, W., Müller, M., & Messner, R. (2012). Teaching methods for modelling problems and students' task-specific enjoyment, value, interest and self-efficacy expectations. *Educational Studies in Mathematics, 79*(2), 215–237. <https://doi.org/10.1007/s10649-011-9341-2>
- Selçuk, E. (2020). *Beceri temelli soru hazırlama etkinliği*. https://bingolodm.meb.gov.tr/meb_iys_dosyalar/2020_11/16112518_beceri_temelli_olcme_ve_d_eYerlendirme.pdf
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Sweller, J. (2012). Cognitive load theory. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 601–605). Springer. <https://doi.org/10.1007/978-1-4419-1428-6>
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.
- Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59–89.
- Taani, O., & Alabidi, S. (2024). ChatGPT in education: benefits and challenges of ChatGPT for mathematics and science teaching practices. *International Journal of Mathematical Education in Science and Technology, 52*11. <https://doi.org/10.1080/0020739X.2024.2357341>
- van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology, 36*(3), 212–218. <https://doi.org/10.1016/J.CEDPSYCH.2010.10.004>
- van Gog, T., Paas, F., & Sweller, J. (2010). Cognitive load theory: Advances in research on worked examples, animations, and cognitive load measurement. *Educational Psychology Review, 22*, 375–378. <https://doi.org/10.1007/s10648-010-9145-4>
- Vural, T. (2020). *An investigation of 8th grade students' statistical literacy through problems with and without context* (Publication No. 638190) [Master's thesis, Middle East Technical University]. Ulusal Tez Merkezi.
- Watson, J. M. (2006). *Statistical literacy at school; Growth and goal*. Lawrence Erlbaum Associates.
- Watts, D. G. . (1991). Why is introductory statistics difficult to learn? And what can we do to make it easier? *The American Statistician, 45*(4), 290–291.
- Wright, K. (2023). *ChatGPT large language model: Everything you need to know*. <https://indatalabs.com/blog/chatgpt-large-language-model>
- Yolcu, A. (2012). *An investigation of eighth grade students' statistical literacy, attitudes towards statistics and their relationship* (Publication No. 321087) [Master's thesis, Middle East Technical University]. Ulusal Tez Merkezi.

- Zhou, R., He, X., Fan, Q., Li, Y., Li, Y., Xiao, X., & Fang, J. (2025). Exploring ChatGPT-facilitated scaffolding in undergraduates' mathematical problem solving. *Journal of Computer Assisted Learning*, 41(4). <https://doi.org/10.1111/jcal.70077>
- Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction*, 4(3), 137–166.
- Zora, L. (2019). *Sekizinci sınıf öğrencilerinin istatistiksel akıl yürütme düzeylerinin incelenmesi* (Publication No. 557444) [Master's thesis, Eskişehir Osmangazi Üniversitesi]. Ulusal Tez Merkezi.

Appendix 1.

ChatGPT-Based Statistics Activity Sheet I

CHATGPT DESTEKLİ İSTATİSTİK ETKİNLİĞİ

Annesi ile markete giden Mehmet, yapacakları tatlı için kaç kilogram ceviz almaları gerektiğini annesine sorar. Annesi ise şu cevabı verir: "1000 gram kabuklu cevizden 500gr ceviz içi çıkar." Mehmet'in annesi, üç farklı alışverişte marketten aldığı kabuklu cevizlerin ve bunlardan çıkan ceviz içlerinin miktarlarını hatırladığı kadarı ile yazarak aşağıdaki tabloyu oluşturur.

Alışverişler	Kabuklu ceviz (gr)	Ceviz içi (gr)
1. alışveriş	800	
2. alışveriş		300
3. alışveriş		

Mehmet annesinin market fişlerinden üç farklı alışverişte aldığı kabuklu ceviz miktarının aritmetik ortalamasının 600 gram olduğunu hesaplamıştır. Buna göre 3. alışverişte aldıkları kabuklu cevizden elde ettikleri ceviz içi miktarı kaç gramdır?

Yönerge: Bu sorunun çözümünde ChatGPT'den yardım almak için aşağıdaki sırayla sunulmuş olan adımları takip ediniz.

1) İlk olarak ChatGPT'de bir oturum açınız. ChatGPT'ye vandaki ifadeyi (istemi) yazınız: "Merhaba, ben altıncı sınıf öğrencisiyim ve denklem kurmayı bilmiyorum. Sana bir soru soracağım ama bu soruyu çözmeni istemiyorum. Senden sadece sorunun nasıl çözüleceğini bana sözel olarak adım adım anlatmanı istiyorum". Şimdi bu ifadeyi göndermek için Enter

tuşuna ya da yukarı yönlü ok işaretine  basınız.

2) Şimdi soruyu Chat GPT'ye yüklemeniz gerekiyor. Bunun için;

- Önce ChatGPT'nin menüsü olan sol alt kenarda bulunan  görseline dokunur
- Sonra telefonunun kamerasını kullanarak sorunun fotoğrafını çekin
- Şimdi fotoğrafın yüklenmesini bekleyin ve yükleme işlemi bittiğinde Enter tuşuna ya da

yukarı yönlü ok işaretine  basın

3) Şimdi karşınızda ChatGPT'nin sorunun nasıl çözüleceğine dair sözel açıklamasını görmekteyiz. Açıklamayı dikkatlice okuyoruz.

4) Sayısal çözümünü görmek için şu ifadeyi yazınız: "Şimdi açıklamaları ile birlikte sorunun çözümünü göstermeni istiyorum, ancak denklem içerecek şekilde olmasın." Şimdi bu ifadeyi

göndermek için Enter tuşuna ya da yukarı yönlü ok işaretine  basınız.

5) Ekranda sorunun ChatGPT tarafından açıklamalar ile birlikte nasıl çözüleceğine dair anlatım görmekteyiz. Açıklamaları ve çözüm basamaklarını dikkatle inceleyiniz.

Appendix 2.

ChatGPT-Based Statistics Activity Sheet II

CHATGPT DESTEKLİ İSTATİSTİK ETKİNLİĞİ

Mine evine davet ettiği arkadaşları için dört çeşit kek yapmaya karar verir. Yapacağı kek çeşidi ne olursa olsun kullandığı un ve yağ oranı değişmemektedir. Her kek çeşidi için belirlediği ölçü şu şekildedir: *Eğer bir su bardağı yağ kullanılıyorsa, 4 su bardağı un kullanılmalıdır.* Mine farklı çeşit kekler için kullandığı un ve yağ miktarlarını aşağıdaki tabloda oluşturmaya çalışmıştır.

Kek çeşitleri	Un (su bardağı)	Yağ (su bardağı)
Kakaolu		2
Havuçlu		1/2
Üzümlü	2	
Sade		

Mine dört çeşit keklerde kullandığı yağ miktarının aritmetik ortalamasını 1 su bardağı olarak hesaplamıştır. Bu durumda sade kek için kaç su bardağı un kullanmıştır?

Yönerge: Bu sorunun çözümünde ChatGPT'den yardım almak için ChatGPT'ye aşağıdaki sırayla sunulmuş olan adımları takip ediniz.

1) İlk olarak ChatGPT'de bir oturum açınız. ChatGPT'ye yandaki ifadeyi (istemi) yazınız: *“Merhaba, ben altıncı sınıf öğrencisiyim ve denklem kurmayı bilmiyorum. Sana bir soru soracağım ama bu soruyu çözmeni istemiyorum. Senden sadece sorunun nasıl çözüleceğini bana sözel olarak adım adım anlatmanı istiyorum”.* Şimdi bu ifadeyi göndermek için Enter tuşuna ya da yukarı yönlü ok işaretine  basınız.

2) Şimdi soruyu Chat GPT'ye yüklemeniz gerekiyor. Bunun için;

- Önce ChatGPT'nin menüsü olan sol alt kenarda bulunan  görseline dokununuz
- Sonra telefonunun kamerasını kullanarak sorunun fotoğrafını çekin
- Şimdi fotoğrafın yüklenmesini bekleyin ve yükleme işlemi bittiğinde Enter tuşuna ya da yukarı yönlü ok işaretine  basın

3) Şimdi karşınızda ChatGPT'nin sorunun nasıl çözüleceğine dair sözel açıklamasını görmekteyiz. Açıklamayı dikkatlice okuyunuz.

4) Sayısal çözümünü görmek için şu ifadeyi yazınız: *“Şimdi açıklamaları ile birlikte sorunun çözümünü göstermeni istiyorum, ancak denklem içerecek şekilde olmasın.”* Şimdi bu ifadeyi göndermek için Enter tuşuna ya da yukarı yönlü ok işaretine  basınız.

5) Ekranda sorunun ChatGPT tarafından açıklamalar ile birlikte nasıl çözüleceğine dair anlatım görmekteyiz. Açıklamaları ve çözüm basamaklarını dikkatle inceleyiniz.

Appendix 3.

ChatGPT's Correct Solution for the First Activity

Harika, bu kez tabloyu ve soruyu net bir şekilde görüyorum. Şimdi sorunun nasıl çözüleceğini adım adım, sözel olarak ve senin anlayacağın şekilde anlatıyorum:

Soru Özeti:

- Mehmet'in annesi, 3 alışverişte kabuklu ceviz almış.
- Tablodaki:
 - 1. alışverişte: 800 gram kabuklu ceviz, ceviz içi bilinmiyor
 - 2. alışverişte: kabuklu ceviz bilinmiyor, 300 gram ceviz içi
 - 3. alışverişte: kabuklu ceviz bilinmiyor, ceviz içi soruluyor
- Anneleri diyor ki:

"Kabuklu cevizlerin ortalama miktar 600 gram."

Çözüm Mantığı:

Şimdi adım adım bu soruyu nasıl düşüneceğimizi açıklayalım:

1. adım: "Ortalama" ne demek?

Toplam değeri, kaç tane sayı varsa ona bölmek demek.

Yani:

$$(1. + 2. + 3. alışverişteki kabuklu cevizlerin toplamı) \div 3 = 600$$

Bunu şöyle düşünebilirsin:

$$\text{Üç alışverişte alınan toplam kabuklu ceviz miktarı} \\ 3 \times 600 = 1800 \text{ gram olmalı.}$$

2. adım: Bildiğimiz kabuklu ceviz miktarlarını topla

- 1. alışverişte: 800 gram kabuklu ceviz alınmış.
- 2. alışverişte ve 3. alışverişte ne kadar alındığı bilinmiyor.

Ama biz toplamda 1800 gram olduğunu öğrendik ya, o zaman şöyle düşünebiliriz:

$$2. \text{ ve } 3. \text{ alışverişte alınan cevizlerin toplamı} = \\ 1800 - 800 = 1000 \text{ gram}$$

3. adım: Her 1000 gram kabuklu cevizden 500 gram ceviz içi çıkıyor (oran: 1/2)

Yani:

Kabuklu ceviz miktarını ikiye bölersen, ceviz içini bulursun.

4. adım: 2. alışverişte 300 gram ceviz içi çıkmış

Az önce bulduk:

2. ve 3. alışverişte toplam 1000 gram kabuklu ceviz alınmış.

Bu 1000 gramın içinden toplam kaç gram ceviz içi çıkmalı?

$$1000 \div 2 = 500 \text{ gram ceviz içi}$$

Peki bu toplam 500 gramın 300 gramı 2. alışverişteyse,

3. alışverişte ne kadar çıkmış olmalı?

$$500 - 300 = 200 \text{ gram ceviz içi}$$

Cevap: 3. alışverişte çıkan ceviz içi miktarı 200 gramdır.

Bu çözümde hiç denklem kullanmadık, sadece mantıkla ilerledik.