

Research

Experimental research on enhancing effectiveness in teaching the nature of seasons

Melike Güzin Semercioğlu¹  · Hüseyin Kalkan² 

Received: 23 June 2024 / Accepted: 30 April 2025

Published online: 19 May 2025

© The Author(s) 2025 

Abstract

In recent years, studies by education researchers indicate that, despite various instructional methods to enhance the conceptual understanding of seasons, the reasons for the formation of seasons and the processes involved are still not fully grasped. This research aims to investigate the fundamental reasons for the difficulties in comprehending the formation of seasons and to develop possible solutions. To achieve this goal, a physical model based on the “Change in the Amount of Energy Falling onto a Unit Surface (CAEFUS)” of parallel beams of sunlight as Earth orbits around the Sun with a tilted axis of approximately 23.5 degrees was developed and applied to participants, and data were collected. These instructional processes were applied to experimental and control groups. The research was conducted on 148 eighth-grade students in a state school in Samsun (the city in Turkey) province. Exploratory factor analysis was used to analyze students’ drawings and separate them into their most prominent elements. The quantitative data analysis, consisting of multiple-choice questions, utilized the SPSS 22.0 statistical program. According to ANOVA Results for SEMAT Achievement Scores, a significant difference was determined between the experimental and control groups ($F = 7.625$, $p < 0.05$). In addition, the control group mean in the pre-test fell below the experimental group mean in the pre-test. However, the mean of the experimental group in the post-test exceeded that of the control group (Pre-test’ $e' = 1.54$, Pre-test’ $c' = 1.82$; Post-test’ $e' = 3.21$, Post-test’ $c' = 2.86$). The data analysis revealed that the CAEFUS-based model significantly impacted the meaningful teaching of the nature of seasons. These results were thoroughly evaluated through a detailed comparison with the literature.

Keywords Nature of seasons · Astronomy education · Physical model · Mental model · Example lesson plan

1 Introduction

This study is grounded in the theory of conceptual change and the constructivist approach to learning. According to Piaget’s theory of cognitive development, individuals construct knowledge through interactions with their environment, progressing through stages of increasing complexity in their understanding of abstract concepts [1]. In science education, this process involves moving from intuitive, everyday understandings to more scientifically accurate models. Models are critical in this process as they help bridge the gap between abstract scientific concepts and students’ existing mental frameworks. This allows for deeper engagement and understanding [2]. The concept of “Change in Energy Amount on Unit Surface” (CAEFUS) aligns with the theory of solar radiation and its distribution, which is fundamental to

✉ Melike Güzin Semercioğlu, guzin.sem@hotmail.com; melikesemercioğlu@gumushane.edu.tr; Hüseyin Kalkan, kalkanh@omu.edu.tr; kalkans@omu.edu.tr | ¹Department of Medical Services and Techniques, Gümüşhane University, Gümüşhane, Turkey. ²Department of Mathematics and Science Education, Ondokuz Mayıs University, Samsun, Turkey.



understanding astronomical phenomena such as seasons. This theory explains how variations in the Earth's orientation and movement relative to the Sun lead to changes in the amount of solar energy received at different latitudes, influencing temperature patterns across seasons [3]. In the context of science education, the use of physical models, as outlined by Bransford and Schwartz supports the idea that concrete representations of abstract phenomena promote deeper understanding by enabling students to visualize complex processes and actively engage with them [2]. In this study, the CAEFUS model serves as a tool to help students understand the mechanisms behind seasonal changes by providing a tangible, interactive representation of solar radiation and its effects on Earth's surface.

Understanding the nature of science involves grasping the mechanisms behind events and the relationships between them [4]. Fundamental astronomy topics, which play a crucial role in science, are gaining importance, supported by growing research and educational strategies [5–7]. While some scientific concepts, like the sunrise, moon phases, and seasons, are observable, the mechanisms driving them, such as celestial movements, remain abstract and not directly observable. The topic of seasons, central to this study, is visible but requires concrete tools for full comprehension. The use of visual aids in science education is widely accepted, although debates persist regarding their role [8]. Early studies demonstrated the effectiveness of graphics in learning, which continues to influence current research [9]. One persistent misconception, highlighted by Van Loon et al. is the belief that Earth's proximity to the Sun causes summer [10].

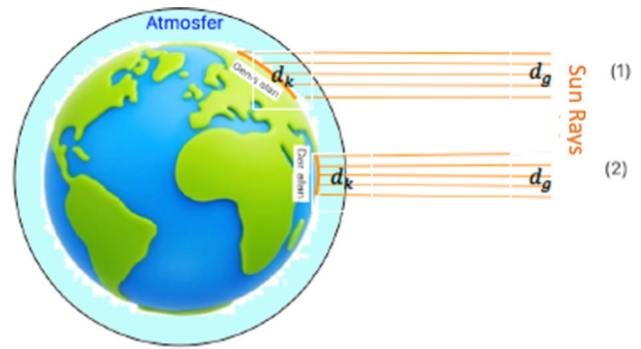
Models play a critical role in science education by concretizing abstract ideas and fostering engagement [11]. However, improperly developed models may lead to misconceptions [12]. Effective use of models can reshape mental models and facilitate learning complex scientific ideas [13, 14], especially in fields like astronomy, where visualization and interaction enhance understanding. Sneider and colleagues reviewed 41 articles on strategies, practices, and challenges in teaching seasons across various countries up to 2011 [15]. The challenges in teaching astronomy concepts arise from their complexity and abstract nature, requiring three-dimensional thinking [16]. Understanding seasons involves aligning temperature patterns with Earth's movements and the geometry of solar radiation distribution [17]. Teaching these concepts effectively can be aided by physical models, which allow students to make scientifically accurate interpretations. Additionally, understanding the geometry of solar radiation distribution and its absorption by the Earth's surface is crucial for comprehending the seasons. Although this academic knowledge can be challenging for middle school students, physical models can effectively teach these concepts and their relationships, enabling accurate scientific interpretations. Many studies affirm that seasons are an excellent tool for teaching scientific principles [17–22]. However, there is a notable lack of studies that specifically address the “Change in Energy Amount on Unit Surface” (CAEFUS) by exploring the “HOW” of understanding this concept, particularly concerning its effects on students' mental models related to Earth's shape, revolution, and axial tilt. This study aims to fill this gap by enhancing 8th-grade students' comprehension of “HOW” seasons occur through the CAEFUS physical model. By doing so, it distinguishes itself from existing research on physical models, addressing a significant gap in the literature. This study provides insights into the effectiveness of physical models in teaching complex astronomical concepts. It offers practical implications for the development of educational tools aimed at improving understanding and rectifying misconceptions in astronomy. The research explores the solutions proposed by students for CAEFUS and addresses the following questions through the use of models:

- Question 1: Have the students' explanations and visual representations of the nature of seasons been influenced by the CAEFUS model? (Is this change significantly in favor of the model?).
- Question 2: How has the CAEFUS model influenced students' explanations and visual representations of the nature of seasons? (Did it take into account the physical events that occurred?).
- Question 3: Which features of the models used most significantly affected students' ability to provide meaningful explanations and visual expressions of the nature of seasons?

1.1 Enhancing understanding of seasonal formation: a conceptual framework

Most studies on teaching the formation of seasons typically address the question, “Why do seasons occur?” The common answers—“axial tilt,” “Earth's movements (rotation and revolution),” and “Earth's distance from and proximity to the Sun”—often fail to address the underlying mechanisms fully. A more relevant question is, “How do seasons occur?” However, research focusing on changes in students' mental models regarding this process based on CAEFUS, which explains the formation of seasons through the angle of sunlight relative to the Earth's surface, remains limited [15, 23]. Students' difficulty understanding the physical processes necessary for CAEFUS is a significant challenge. Figure 1 illustrates the variation in solar radiation distribution on the Earth's surface. When sunlight is perpendicular to the surface, the area it covers is smaller, resulting in higher energy concentration. Conversely, at an oblique angle, sunlight spreads

Fig. 1 The physics of solar incidence angle



over a larger area, leading to lower energy concentration. These variations are critical to understanding the formation of seasons but can be abstract and difficult for students to visualize.

Figure 1 illustrates the variation in the distribution of incoming solar radiation across the Earth's surface. In position (2), where sunlight is perpendicular to the Earth's surface, the area covered by sunlight is relatively more minor. Conversely, in position (1), where sunlight strikes the surface at an angle, the area over which the sunlight is spread is significantly larger. These variations are crucial for understanding the concept of seasons. Terms such as "perpendicular sunlight" and "oblique sunlight" clearly represent how these variations influence seasonal changes.

Figure 2 presents the mathematical analysis of the thermal distribution effect on the Earth's surface resulting from the natural phenomenon depicted in Fig. 1. The relationship between the angle of incidence of rays and the spread occurring on the Earth's surface is as follows:

$$d_k = \frac{d_g}{\sin(\theta)} \quad (1)$$

In Eq. (1), d_k represents the area where solar rays spread on the Earth's surface, d_g represents the area that solar rays have before encountering any surface, and θ is the angle of incidence as shown in (a), when $\theta = 30^\circ$ the area where solar rays spread is more expansive, while when $\theta = 90^\circ$ the area where rays spread is narrower. Therefore, the amount of energy per unit area incident on the Earth's surface is inversely proportional to the area over which solar light spreads. For example, assuming that the energy of a beam of light is 10 J, and it spreads over surfaces of 10 units and 20 units separately, in the first case, 1 J (10/10) of energy falls on a unit area, while in the second case, 0.5 J (10/20) of energy falls on a unit area. The same situation can be applied to (b). Parallel incoming solar rays will spread over different areas at different angles depending on the surface shape they encounter. Accordingly, the energy transfer per unit area will vary inversely with the width of the surface.

This equation represents the area where solar rays spread on the Earth's surface, the area that solar rays have before encountering any surface, and the angle of incidence. For example, the area where solar rays spread is wider than when. Consequently, the amount of energy per unit area incident on the Earth's surface is inversely proportional to the area over which solar light spreads.

All factors contributing to the formation of seasons, as illustrated in Fig. 3, represent a scientifically grounded framework for understanding the interplay between physical and astronomical phenomena. The CAEFUS model emphasizes the angle of the Sun's rays relative to the Earth's surface and integrates multiple scientific principles crucial for students to grasp, including:

Fig. 2 Heat transfer according to the angle of incidence of rays

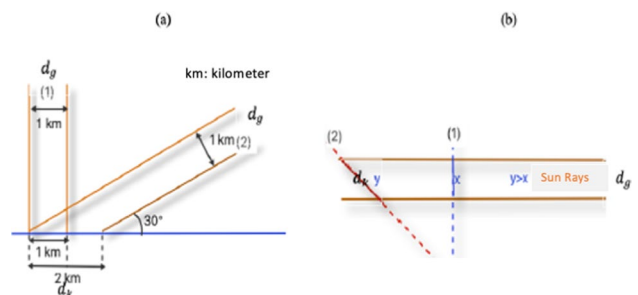
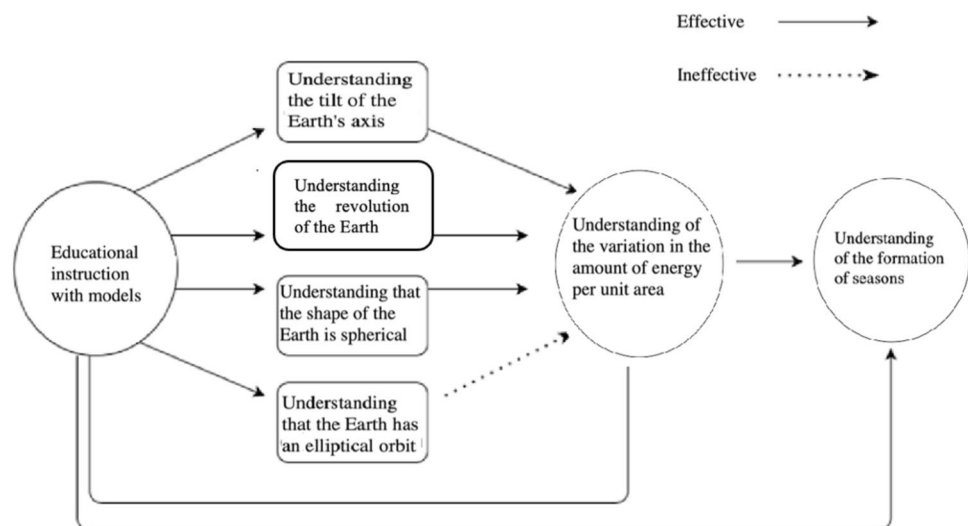


Fig. 3 Factors affecting the amount of energy falling on unit surface



The CAEFUS model integrates physical and astronomical concepts to understand seasonal changes comprehensively. Figure 3 outlines the framework, emphasizing the following key factors:

- *Axial Tilt:* The Earth's 23.5° axial tilt results in differential solar energy distribution across hemispheres, affecting sunlight intensity and duration.
- *Earth's Revolution:* The elliptical orbit causes varying sunlight angles across regions, underpinning seasonal cycles.
- *Earth's Spherical Shape:* The curvature of the Earth influences sunlight angles by latitude, concentrating energy near the equator and dispersing it toward the poles.

The core concept—variation in energy per unit area—serves as the foundation for interpreting solar energy distribution and its role in seasonal patterns. The CAEFUS framework encourages students to refine their mental models through physical and visual representations, addressing common misconceptions and fostering more profound understanding.

This study addresses challenges in teaching the formation of seasons by exploring “HOW” physical phenomena contribute to seasonal changes using the CAEFUS model. By analyzing students’ pre- and post-instruction explanations, the study evaluates their alignment with scientific accuracy. Grounding the teaching process in these principles ensures that students move beyond rote memorization to understand the mechanisms behind seasonal changes comprehensively.

In conclusion, the CAEFUS model provides educators with a structured approach to teaching the formation of seasons, integrating scientific principles and addressing misconceptions. This method promotes the development of accurate mental models and enhances students’ cognitive engagement with the subject.

2 Method

Morse’s “Simultaneous Quantitative Dominant Mixed-Method Research Method” was employed in this study. In this method, qualitative and quantitative data are collected simultaneously and analyzed separately, and the results are interpreted in congruent and incongruent aspects. This approach necessitates using a mixed-method (MM) approach [24]. Furthermore, this study utilized a quasi-experimental method with two samples [25]. The sample consisted of 148 (72 experimental group + 76 control group) 8 th-grade students attending a public school affiliated with the Ministry of National Education in Samsun province. The students included in the experimental and control groups were determined using a simple systematic sampling method. The students were assigned to the experimental and control groups one by one based on an alphabetical order of their last names. This is a non-random assignment, as the students were not randomly assigned to the groups. Therefore, the study is classified as quasi-experimental rather than entirely experimental. Pre-test and post-test measurements were administered before and after instruction, as well as during the retention test. Both qualitative and quantitative measurements were utilized during the implementation phase of the study. In the qualitative dimension of the research, in-depth analyses were conducted using an open-ended question scale to explore context-dependent themes and identify students’ mental models in detail.

The study's main theme is to assist in structuring mental models related to CAEFUS. Changes in mental models can be determined by measuring the development of practical problem-solving skills. Therefore, the data collection process used a mixed-method (MM) approach. The stages of the research process are outlined in Fig. 4. The teaching of seasons is part of the science education curriculum for 8th-grade students. In this study, all 8th-grade students were combined and alphabetically sorted by their last names. They were then distributed one by one into the control and experimental groups. 8th-grade students formed the experimental and control groups from the same school. This aimed to ensure that both groups were exposed to the same education from the same teachers. I, as the researcher for both groups, implemented the study. The instructional process of the study lasted a total of 8 lesson hours, with each group receiving $2 + 2 + 2 + 2$ h of instruction. Both groups received 8 h of education related to the teaching of seasons. Pre-tests were conducted before instruction, post-tests after instruction, and a retention test was administered 8 weeks after the post-test. To determine the internal consistency of the Solar and Earth Movements Achievement Test (SEMAT), the KR-20 internal consistency coefficient was calculated. The item analyses and KR-20 internal consistency coefficients of the pilot and final SEMAT are presented in Table 1. The KR-20 formula is based on the assumption that each item in the test measures the same variable and that the test measures the same characteristic [26]. The reliability coefficient ranges from 0.00 to 1.00, with values between 0.80 and 1.00 indicating high reliability, between 0.60 and 0.80 indicating quite reliable, and below 0.60 indicating low or very low reliability [27]. The KR-20 internal consistency coefficient for the final SEMAT was calculated as 0.77. Since the lower limit of the reliability coefficient for measurement tools in research is 0.70, this study provides an acceptable level of reliability. The analysis results confirm that SEMAT is a valid and reliable test for measuring teacher success in astronomical topics.

The initial motivation for this study arose from the observation that, while students understand the reasons for the formation of seasons, awareness of the underlying physical processes is often lacking. For example, although they recognize that the Earth's revolution around the Sun affects the formation of seasons, they do not fully grasp the physical consequences of this movement and its impact on seasonal changes. As illustrated in Fig. 4, this realization led us to develop a research plan and a corresponding physical model to address this gap. A comprehensive literature review was conducted at the outset. While being aware of similar studies in the field, the focus was explicitly on those addressing physical processes or identifying similar gaps in understanding [23, 28–34]. Preliminary qualitative and quantitative scales were developed to align with the study's objectives: the Solar and Earth Movements Achievement Test (SEMAT) and the Solar and Earth Movements Open-Ended Questionnaire (SEMONEQ). Expert feedback was sought during the development of SEMONEQ to ensure the questions remained focused and concise. Pilot studies validated these scales, and the reliability and validity of SEMAT were confirmed using SPSS statistical analysis. The pilot study conducted during the development of the scales used in the research was applied to a group different from the final sample group before the actual research. Although this research forms part of a more extensive study, pilot data have not been included here to avoid lengthening the article but are available upon request.

Simultaneously, we developed a draft physical model with expert input while formulating the research questions. This model was applied multiple times to various groups to refine and address shortcomings. The experimental group implemented the finalized model, while the control group followed the standard science curriculum. The instructional process and data collection were carried out concurrently. Both groups were administered a pre-test before instruction, a post-test after instruction, and a retention test eight weeks later. The collected data were analyzed through content analysis and SPSS, with findings compared to the conceptual framework and discussed in the conclusion. Participants were not actively involved in the model's development process. The researcher independently designed and developed the model, which was then implemented as part of the instructional phase (intervention). The intervention involved applying the finalized model to an experimental group while a control group followed the standard science curriculum. The study's sequence of activities is outlined below:

Model Development: The researcher developed a physical model based on the CAEFUS concept, ensuring it aligned with the study's objectives. This phase was conducted independently and did not involve any participants.

Pre-Test Administration: Both experimental and control groups were administered a pre-test (SEMAT and SEMONEQ) to assess their baseline understanding of the seasons.

Instructional Phase (Intervention): The experimental group interacted with the CAEFUS-based model during instruction, while the control group followed the standard science curriculum. Instruction focused on understanding the physical phenomena underlying seasonal changes, such as axial tilt and solar energy distribution.

Post-Test Administration: After the intervention, both groups were administered a post-test to measure the immediate impact of the instructional methods.

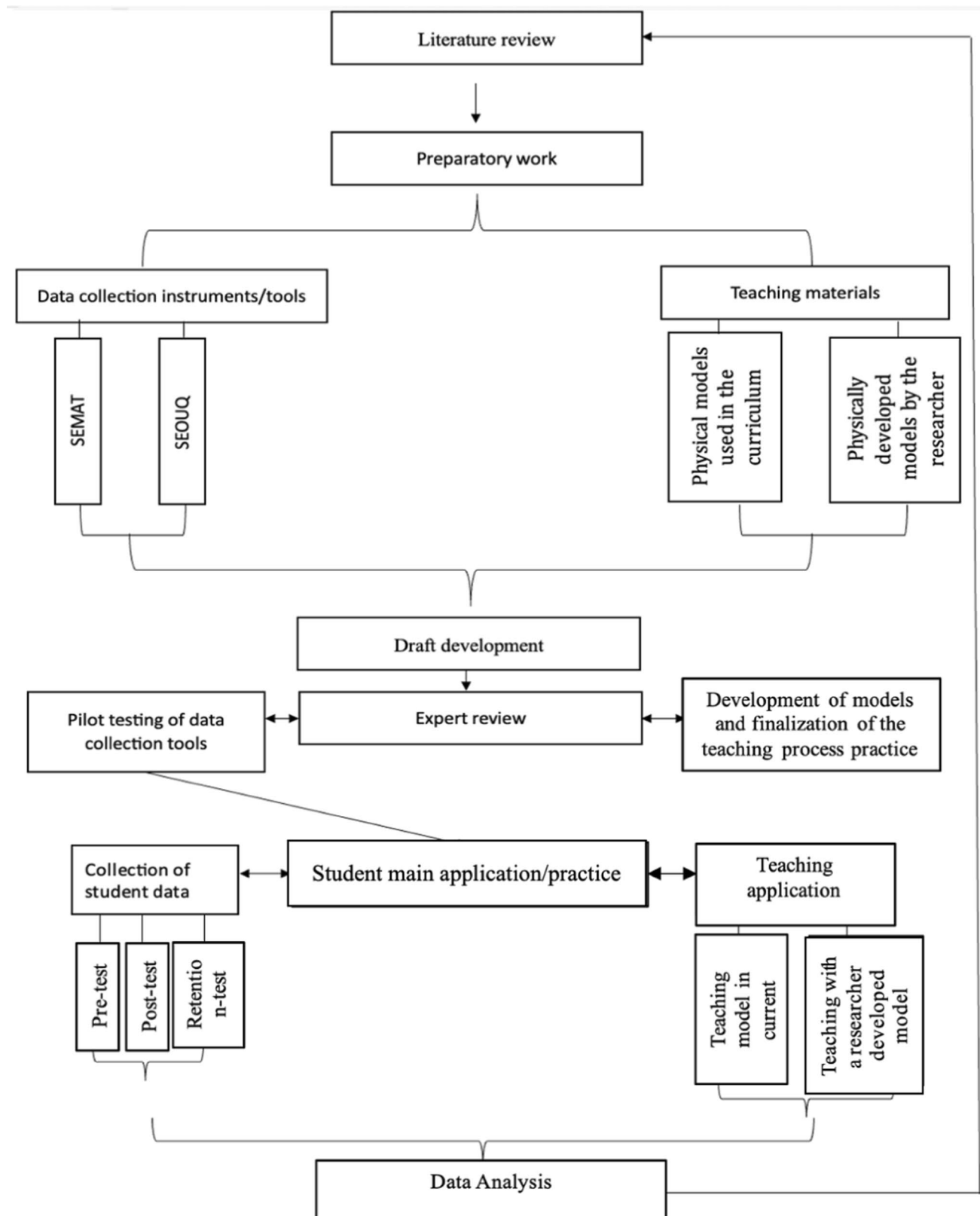


Fig. 4 Stages of the research process

Table 1 Test statistics obtained from item analysis

Item count	Average Difficulty (pj)	Average discriminability (rjx)	KR-20
Pilot	0.58	0.39	0.24
Final	0.47	0.56	0.77

Retention Test: Eight weeks later, a retention test was conducted to evaluate the long-term retention of the concepts taught.

2.1 Data collection and analysis

Quantitative data were collected through SEMAT, a multiple-choice test administered as a pre-test, post-test, and retention test. Qualitative data were obtained using SEMOEQ, an open-ended questionnaire that allowed students to explain and visually express their understanding of seasons. Quantitative data were analyzed using ANOVA and repeated measures analysis with SPSS to compare group performance over time. Qualitative data were coded and categorized as scientifically correct, partially correct, or incorrect to identify student conceptual understanding changes. Additionally, one of the quantitative analysis methods, “content analysis,” was used in the qualitative data to examine and interpret the data systematically.

2.2 Stages of developing the seasons model

During the development of the seasons model, the concept of “the amount of energy falling onto a unit surface depending on the angle of the Sun’s rays with the Earth’s surface” was planned to be taught. The model was constructed according to this plan.

- Scientific information regarding the Sun and the Earth was collected, and the scaling in the model was calculated. Information regarding the distance between the Earth and the Sun is provided in Table 2.
- Different positions (significant dates) were determined on the model to illustrate the Earth’s orbit around the Sun, and a fixed but practical model was chosen instead of a movable one.
- Since it was impossible to create the actual sizes of the Sun and the Earth within the same model, three powerful light-emitting sources were used instead of the Sun.
- Four Pilates balls with a diameter of 85 cm, representing the Earth’s four seasons, were drawn with latitudes and longitudes, and fixed systems were created to determine Turkey’s position on the Earth. The models used in the research are given in the appendix at the end of the article (Figs. 10 and 11).
- The axes of the Earth were positioned to indicate the North Star and were fixed on supports at a height of one meter. Essential dates were written on the supports, and the positions and approximate distances of the Earth on these crucial dates were calculated.
- Three identical (with a capacity of 1000 m) vertical lights were used in the same direction through the centers of the four Earth models (considering the distances to the Sun). There was a gap of 13 cm between the lights.
- The lights were positioned so that on March 21 and September 23, they would fall on the Tropic of Cancer at the top, the Equator in the middle, and the Tropic of Capricorn at the bottom.

Table 2 Actual and approximate distances between the sun and the earth

Date	Actual distance (approximately)	Scaled distance between sun and earth
March 21	between 147,000,000 and 152,000,000 km	139 cm
June 21	between 147,000,000 and 152,000,000 km	140 cm
July 4 (Aphelion)	152,000,000 km	142 cm
September 23	between 147,000,000 and 152,000,000 km	139 cm
December 21	between 147,000,000 and 152,000,000 km	138 cm
January 3 (Perihelion))	147,000,000 km	137.3 cm

Based on this information, Fig. 5 presents the principles underlying the model development process and the steps related to the model's construction.

To make this study more understandable, Fig. 5 can be explained as follows;

In this study, four Pilates balls and three light sources (flashlights) were used to develop a model of the seasons. The flashlights were fixed to a vertical stand, each positioned 13 cm apart. The stand was placed at the center of the model. Each Pilates ball was marked with the equator and other parallel circles, and a specific region on Earth (Turkey's location in this study) was indicated. Straws were placed at the Pilates balls' north and south pole points to represent the Earth's axis. The Earth's axial tilt of 23.5 degrees was considered in the model, and the Pilates balls were positioned according to the straw indicating the North Pole.

In the first stage of the model (Fig. 6A), the Earth was held in a fixed position, but the four Pilates balls were placed in different locations (March 21, June 21, September 23, and December 21 equinoxes). The three flashlights on the wheeled stand, representing the Sun, were positioned at the model's center and served as a reference point to show the energy received by different regions on Earth. The Earth was rotated around the Sun while maintaining an axial tilt of 23.5 degrees. In this scenario, it was observed that each region received the same amount of energy per unit area within itself, and the same season occurred consistently in different regions of the world. Using Turkey as an example, it was emphasized that the four seasons can only be observed in certain areas.

The axial tilt was eliminated in the model's second stage (Fig. 6B), and the poles were adjusted to indicate 270 degrees and 90 degrees. The Earth was rotated around the Sun at this stage, and the same energy distribution was observed. It was stated that, in the absence of axial tilt, the same amount of energy per unit area would fall on different locations on Earth throughout the year. This scenario attempted to demonstrate that no region would experience different seasons and that some areas of the Earth would experience summer throughout the year. In contrast, others would experience winter or spring.

Finally, an additional flat board was used to explain the Earth's sphericity (Fig. 6C) and its effect on the seasons. By reflecting the three flashlights representing the Sun onto the flat board, it was demonstrated that the energy falling per unit area was the same across all regions. When the flat board was then lifted, it was observed that the rays from the upper and lower flashlights spread over a large area due to the Earth's spherical shape. This showed that the energy per unit area was lower in areas where the rays were spread and higher in areas where the rays, such as at the equator, were less spread. When the three factors were combined, it was explained how the energy falling per unit area varies across different coordinates of the Earth throughout the year and, as a result, how seasons are experienced in other regions.

3 Data collection tools

Various data collection tools were carefully developed and validated to effectively investigate students' understanding of solar and Earth movements.

- a) Solar and Earth Movements Achievement Test (SEMAT),
- b) Solar and Earth Movements Open-Ended Questionnaire (SEMÖEQ).

Firstly, questions were developed considering the student levels and the related learning outcomes. Information obtained from national and international sources was utilized while creating the test. Concepts related to "Earth and the Universe" were identified. It was planned to write at least three questions designed differently but aimed at assessing the same concept. The first draft of SEMAT was prepared to consist of 13 questions. Expert opinions were sought to ensure the content validity of SEMAT. Experts were informed about the test's purpose and were asked to review and evaluate it accordingly. Based on the experts' evaluations, necessary revisions were made, and the pilot implementation phase was initiated. The pilot implementation of the draft SEMAT was conducted in five middle schools in the Gümüşhane city center. The pilot study of the draft achievement test was carried out with 65 students. It was determined that the students completed the test in approximately 25–30 min. A normality test was conducted on the pilot data, and it was found that the data were normally distributed ($p > 0.05$).

The items of the astronomy achievement test were analyzed to determine their validity and reliability. Correct answers were given 1 point in the scoring process, while incorrect answers, unanswered questions, and those with more than one marked option were scored as 0 points. The total score each student received from the test was calculated. An item

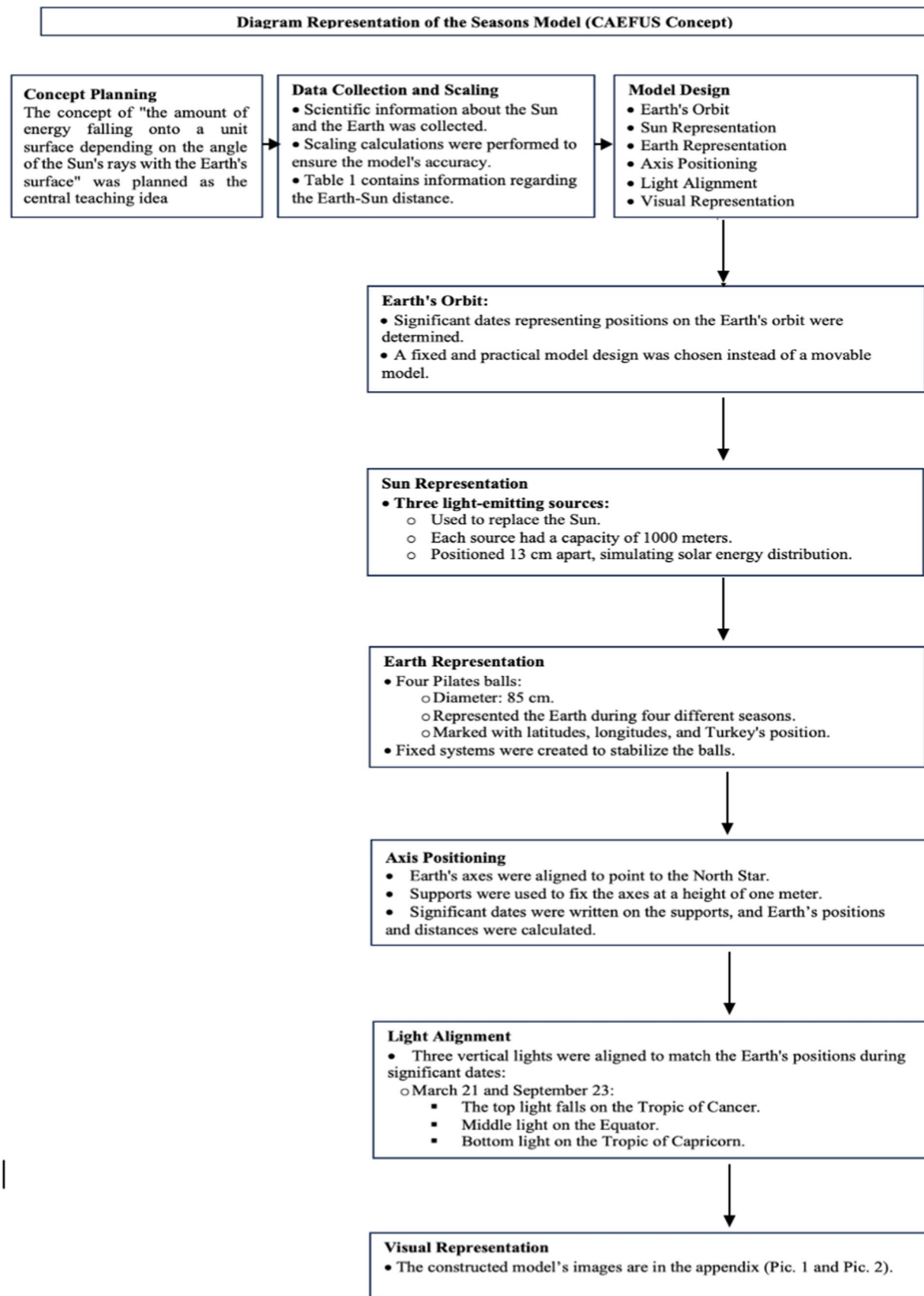


Fig. 5 Structured development of the seasons model based on the CAEFUS concept

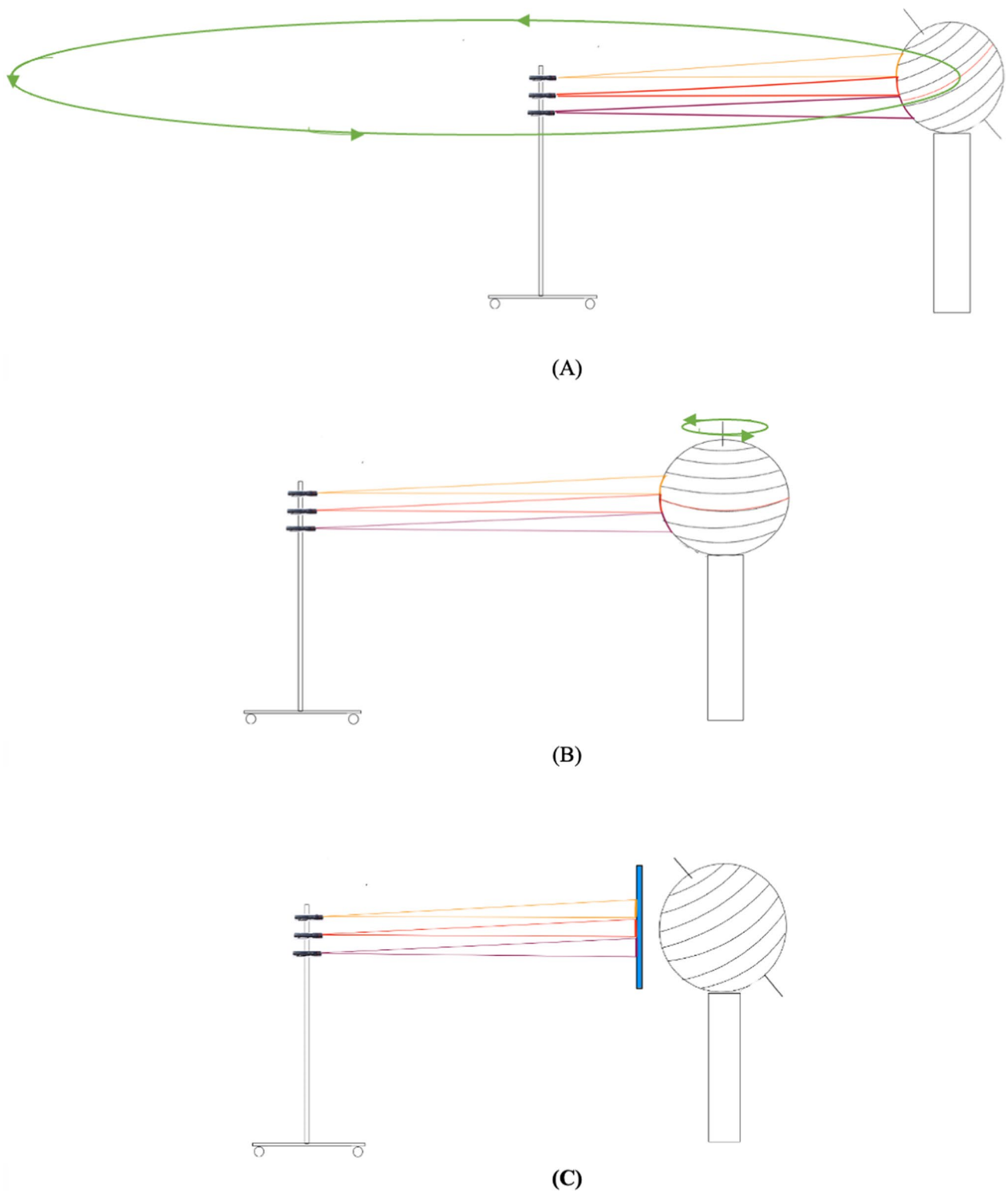


Fig. 6 The application steps of CAEFUS in the formation of seasons: **a** Effect of Earth's orbit around the Sun, **b** Effect of axial tilt, **c** Effect of a flat or spherical surface using a panel

analysis was conducted based on the responses provided for each item. After scoring, the test scores were ranked from highest to lowest. Groups were formed by taking the top 27% of the scores as the upper group and the bottom 27% as the lower group, and item analysis was performed. The item discrimination and difficulty indices were considered during the

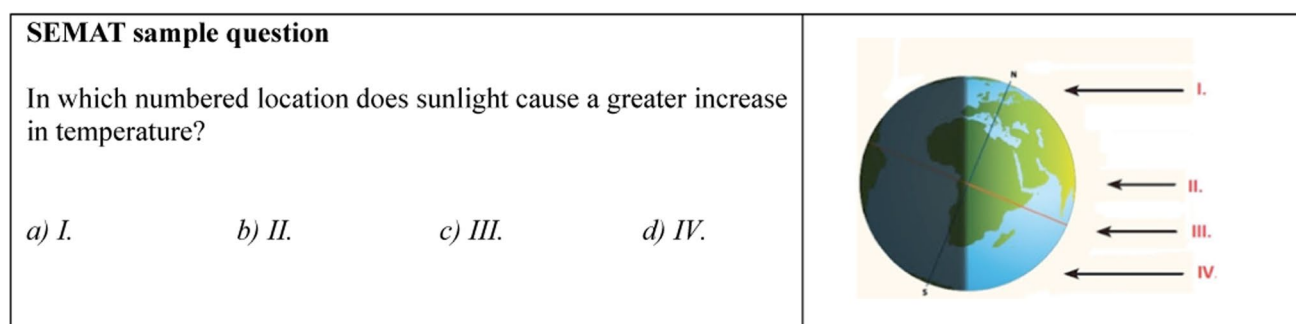


Fig. 7 Example question from SEMAT

item analysis. Questions with low discrimination and unsuitable difficulty indices were either revised or eliminated if they could not be revised. It was deemed appropriate to ask five questions in the final test. The final test's reliability coefficient (pj)¹ was found to be 0.64, and the validity coefficient (r_{jx})² was found to be 0.68. Accordingly, although the scale is easy, its discrimination power is excellent [35–37]. An example question from SEMAT used in the study is as follows (Fig. 7):

This study was conducted as part of a broader astronomy research project. The open-ended question form was administered before SEMAT to ensure that students were not influenced by or prompted by the multiple-choice options in SEMAT. The SEMOEQ, initially consisting of 10 questions, was reduced to 5 based on expert feedback, one of which focuses on the concept of “seasons.” The questions were developed with input from five experts. The open-ended question addresses the learning outcomes of “seasons” in the 8th-grade science curriculum, ensuring content validity. Only the necessary questions were included to minimize distractions and maintain focus on the target learning outcomes. The time allocated for answering the question was set at 15 min. The open-ended question regarding the “seasons” learning outcome posed to students in the study is as follows: “Explain the factors influencing the occurrence of seasons and their reasons. (You may include drawings).”

4 Data analysis

In this study, data were collected using the SEMAT and SEMOEQ scales. These data sets were collected simultaneously, analyzed separately, and interpreted together.

4.1 Analysis of quantitative data (SEMAT)

In this study, data were collected using the SEMAT and SEMOEQ scales. These data sets were collected simultaneously, analyzed separately, and interpreted together. The analysis of SEMAT was carried out as follows: The data obtained from the implementation of SEMAT were analyzed using the SPSS 22.0 statistical software. The test applied in the study was administered three times: pre-test, post-test, and retention test. Before the implementation, classes were randomly selected and classified into control and experimental groups. The same scale was applied to both groups before and after the implementation, with equal time intervals and conditions. The retention test is typically conducted between 1 week and 6 months after the post-test in experimental studies. This duration provides a sufficient time frame to measure how long students retain the information they have learned. However, this period may vary

¹ The item difficulty index (pj) represents the ratio of the number of students answering the item correctly in both the upper and lower groups to the total number of students in those groups. This index ranges from 0.0 to 1.0, and as the value approaches zero, the difficulty level of the item increases, while as it approaches one, the difficulty level decreases. Typically, a difficulty index of around 0.50 is desirable in achievement tests, as it indicates a moderate level of difficulty. The item difficulty index can be classified as follows: 0.00–0.19 is very difficult, 0.20–0.34 is difficult, 0.35–0.64 is moderately difficult, 0.65–0.79 is easy, and 0.80–1.00 is very easy [35, 36].

² The item discrimination index (r_{jx}) indicates the degree to which an item differentiates between high-achieving and low-achieving students. This index ranges from -1 to +1, with values closer to zero indicating low discrimination, values closer to +1 indicating high discrimination, and values closer to -1 indicating high discrimination in the opposite direction. The item discrimination index can be classified as follows: 0.19 and below is very weak, 0.20–0.29 requires improvement, 0.30–0.39 is good, and 0.40 and above is excellent [34–36].

depending on the purpose, topic, and characteristics of the participants in the study. The retention test is usually planned to take place over a longer time frame, as its purpose is to observe how long the learned information is retained. In this study, the retention test was conducted 8 weeks after the post-test. This duration was selected as an adequate time frame to assess how long and to what extent students retained the information they learned. The 8-week period aims to provide a more reliable indicator of the long-term retention process of the information [34]. The pre-test assessed the students' prior knowledge of the concept under investigation before the teaching process. The post-test assessed whether the teaching had achieved its intended outcomes after the intervention. Eight weeks after the post-test, the same test was administered again as the retention test. Descriptive statistical analyses (normality tests, validity, and reliability analyses) were conducted on the data obtained from SEMAT. Throughout the analysis process, the normal distribution of the data was examined to determine whether parametric or non-parametric analysis techniques should be used. The sample size should exceed 30 to perform parametric analysis, and the data must follow a normal distribution. In this study, the number of participants in the experimental and control groups exceeded 30, satisfying the first assumption. Another criterion, the Kolmogorov–Smirnov normality test, indicated a p-value greater than 0.05, confirming the normal distribution of the data.

Before conducting parametric analyses, it is essential to check whether the assumptions of the ANOVA technique, which is planned for use in the analysis, are met [38]. The assumptions of the ANOVA technique are as follows:

- Measurements of the dependent variable are at least interval-level.
- The scores of the dependent variable are normally distributed within each subgroup.
- The variances of the scores obtained in the groups are equal.
- The covariances of the groups' pairwise combinations are equal.
- The difference score calculated for any subject is independent of the difference score calculated for other subjects.

SEMAT is an interval-level scale, and the dependent variable scores are normally distributed in each subgroup. Thus, the first two assumptions are met. To determine whether the variances of the scores obtained in the experimental and control groups are equal, Levene's Test was applied to the pre-test, post-test, and retention test scores of the groups to assess variance homogeneity. The results indicated no significant difference between the variances.

Repeated measures designs differ from classical variance analysis. In repeated measures, observations obtained from the same subject over consecutive short time intervals may be more highly correlated than those obtained over longer intervals. In repeated measures data, this correlation structure is referred to as the variance–covariance structure of the data. A covariance model that includes the actual covariance structure is needed to obtain accurate results from the analysis of repeated measures data. Covariance analysis (ANCOVA) is typically used in pre-test/post-test control group designs to test whether there is a significant difference between the post-test scores of the experimental and control groups. The pre-test scores are defined as covariates [38]. The logic of covariance analysis is to remove the variance in the dependent variable attributable to the covariate and then determine whether the remaining variance in the dependent variable is attributable to the independent variable. ANCOVA is the most common method for controlling a variable in analysis, i.e., eliminating its effect [39]. If the groups are unequal, ANCOVA can also be used to equalize them. For example, in a study comparing different teaching methods using non-randomly selected students, there may initially be a difference in intelligence between the groups. If the groups are believed to be unequal, ANCOVA can be used to equalize them. In this way, the effect of intelligence is removed before comparing group means on the dependent variable. When randomly assigned groups, ANCOVA is an excellent method for comparing differences between two groups [40]. The assumption of equality of covariance matrices between groups, the following assumption in the study's sample group, was analyzed using Box's M test. If the p-value obtained from Box's test is greater than 0.05, covariance equality can be assumed.

The covariance equality of the groups concerning SEMAT scores was tested, and it was determined that the covariances were equal ($F_{6-160156.255} = 4.304$, $p > 0.05$). Thus, the fourth assumption was also met. Moreover, the difference score calculated for any subject is independent of the difference score calculated for other subjects, meeting the final assumption. In this study, the mixed-design ANOVA technique was selected to assess the astronomy achievement of students in the experimental and control groups. The comparison of the experimental and control groups involved "independent samples," while the repeated application of the pre-test, post-test, and retention tests involved "related samples." Because this method encompasses the measurement of different groups and the repeated measurement of the same groups at other times, a mixed-design variance analysis method was used.

4.2 Analysis of qualitative data (SEMÖEQ)

The answers given by students to open-ended questions were categorized as scientifically correct, scientifically partially correct, and scientifically incorrect. Students who made an entirely correct explanation by drawing correctly on the diagram were coded as scientifically correct; students who made a partially correct explanation by providing incomplete explanations on the drawing or only the explanation were coded as scientifically partially correct; students who made both the drawing and the explanation wrong were coded as scientifically incorrect. The qualitative data collected in this study dimension were analyzed using the content analysis method, and a quantitative coding approach was applied. The percentages and frequencies of the obtained data were determined. Changes in students' mental models, defined by their drawings for 3 tests before and after the application, were analyzed.

The SEMÖEQ analysis applied to the students was analyzed using the content analysis method. In this process, the responses and drawings to the question "Explain how seasons are formed by drawing" were examined, and codes and categories were created for the expressions and drawings. First, while reviewing the pre-test data, the papers were read individually, and no codes or categories were created during this process. After gaining a general understanding, building codes and categories for the expressions began. While determining codes, two researchers independently worked on the same dataset, dividing it in half. Then, a consensus was reached regarding whether the codes were consistent and whether the agreed-upon codes were used as they were. Discussions were held to resolve discrepancies and reach common ground. Open-ended question forms were re-read through the codes. The expressions were classified based on content similarity to determine the mental models for the responses, and names were given to each classification to define the main categories [41]. The findings were presented in tables under the findings section. When creating the tables, category diversity was considered, and repetitive definitions were omitted. Furthermore, expressions that addressed multiple categories were evaluated in various categories. Therefore, it can be expected that the number of expressions in the tables may differ from the number of assessed forms. To ensure reliability in the data analysis of qualitative research, another expert researcher in the field independently repeated the data analysis process, and the consistency of the codes and categories created by the researcher was checked. In the analysis, consistent codings were indicated as "Agreement," while inconsistent codings were noted as "Disagreement." The reliability calculation obtained from these codings was performed using the following formula [42].

$$\frac{\text{Agreement}}{\text{Agreement} + \text{Disagreement}} \times 100 \quad (2)$$

Two separate researchers analyzed the open-ended questions and created 9 different categories for each. Both researchers created 10 coded categories for each question. Eight were common, while one category differed from the first researcher and one from the second researcher.

$$\frac{9}{9 + (1 + 1)} \times 100 = \%81.8 \quad (3)$$

Based on this result, the consistency percentage of the data analysis was determined to be 81.8%. According to Şimşek and Yıldırım [43] a consistency percentage of 70% indicates that the coding is reliable. Examples of categories of disagreement between the two researchers are provided below. In the question asking for an explanation of how seasons are formed, one researcher created a category called "Change in the angle of solar radiation," while the other made a category called "Change in the angle at which solar radiation reaches the Earth." Both categories were merged to form a single category, "Change in the angle of solar radiation relative to the Earth's surface."

5 Findings and interpretation

This section focuses on the findings resulting from the analyses that were conducted. The findings are organized and presented to address the sub-objectives. Visual elements such as tables, figures, and graphs are used in presenting.

5.1 Findings related to SEMAT

At this stage, the pre-test results of the experimental and control groups were evaluated using one-way ANOVA for unrelated samples. Table 3 presents descriptive statistics for the pre-test results of students' astronomy achievements.

Table 3 highlights the experimental and control groups' remarkably close pre-test mean scores. This led to the conduct of one-way ANOVA to assess the score differences between the experimental and control groups, which were formed according to students' grade levels. The results of the analysis are presented in detail in Table 4.

According to Table 4, there is no significant differentiation before experimental application ($F_{(4,74)} = 1.741, p > 0.05$). This result indicates that the groups' achievements related to the concepts and events under study were equivalent before the intervention.

To assess whether there is a significant difference in the achievement levels between the experimental and control groups, one-way analysis of variance (ANOVA) technique was used for the measurements. Descriptive analysis A of the experimental group students' SEMAT pre-test, post-test, and retention test results are presented in Table 5. This table will assist us in visually understanding the analysis results.

According to Table 5, the experimental group showed an increase in the average scores in the SEMAT analyses on the post-test and a greater increase in the retention test. A total of five questions were asked to the eighth-grade students. The average score was 1.52 on the pre-test, increased to 3.21 on the post-test, and rose to 3.76 on the retention test. In the control group, the average score was 1.82 on the pre-test, increased to 2.87 on the post-test, and rose to 3.30 on the retention test. Given that five questions were asked, the numerically small values carry significant meanings. To determine whether the SEMAT achievement scores of the experimental and control groups differed significantly, a one-way repeated measures analysis of variance was conducted for each group. The data for the one-way analysis of variance are presented in Table 6.

According to Fig. 8, the average scores for pre-tests, post-tests, and retention have increased in the experimental group of students. Five questions were asked. The average score was 1.52 on the pre-test, which increased to 3.21 on the post-test, and the average rose to 3.76 on the retention test. Findings regarding the statistical significance of the data obtained from the tests are provided in Table 6.

The pre-test results of the control group students showed an average score of 1.82. This average increased to 2.87 in the post-test. The retention test average was 3.30.

A one-way repeated measures ANOVA was conducted to determine whether there was a significant difference in the SEMAT achievement scores of the experimental group students. The data for this ANOVA are provided in Table 6.

Table 3 Descriptive statistics for SEMAT

Group	N	Arithmetic mean	SD	Skewness	Kurtosis
Experimental	72	1.52	0.905	0.555	0.252
Control	76	1.82	1.016	0.256	− 0.402

Table 4 One-way ANOVA results for SEMAT pre-test scores of experimental and control group students by class

Source of variance	SS	df	MS	F	p
Between groups	1.741	4	0.435	0.517	0.724
Within groups	58.979	70	0.843		
Total	60.720	74			

Table 5 Descriptive statistics for SEMAT pre-test, post-test, and retention test results of experimental group students

Group	Test	N	Arithmetic mean	SD	Skewness	Kurtosis
Experimental	Pre-test	75	1.52	0.90	0.562	0.264
	Post-test	76	3.21	1.23	− 0.073	− 0.851
	Retention	77	3.76	0.98	− 0.618	0.281
Retention	Retention	78	1.82	1.01	0.238	− 0.439
	Retention	80	2.87	1.17	− 0.321	− 0.278
	Retention	79	3.30	1.13	− 0.837	1.414

Table 6 One-way ANOVA results for students' SEMAT achievement scores

Group	Source of variance	SS	df	MS	F	p	η^2
Experimental	Pre-post-retention test	201.400	2	100.700	91.396	0.000	0.449
	Error	246.802	224	1.102			
	Total	448.203	226				
Control	Pre-post-retention test	90.891	2	45.445	36.809	0.000	0.240
	Error	287.664	233	1.235			
	Total	378.555	235				

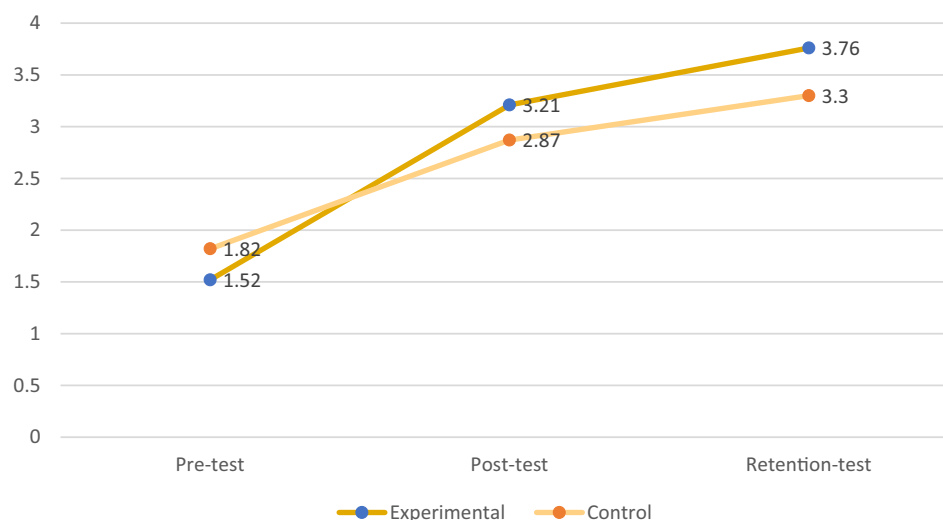
Fig. 8 Changes in SEMAT scores by group throughout the process

Table 6 shows a significant difference between the pre-test and post-test scores of the experimental group ($F = 91.396$, $p < 0.05$). These results indicate that the teaching model used was effective in instruction and led to a significant differentiation. They were additionally, based on the eta squared (η^2) value, it can be said that the teaching with the Solar-Earth models developed within the scope of the research had a considerable impact on students' SEMAT achievements ($\eta^2 = 0.449$).

In the control group, there is also a significant difference between the pre-test and post-test scores ($F = 45.445$, $p < 0.05$). These results also indicate that the teaching model used was effective in instruction and led to a significant differentiation. Moreover, according to the eta squared (η^2) value, the impact of the teaching method used on students' SEMAT achievements can be considered moderate ($\eta^2 = 0.240$).

Eta squared is a statistical value that measures the effect of an independent variable on the dependent variable. It expresses the percentage of the total variance in the dependent variable explained by the independent variable and usually falls within the range of 0.00 to 1.00. Eta squared helps us determine the effect size for each main effect, interaction, and error in ANOVA analyses. These statistics allow us to assess the magnitude of the independent variable's impact on the dependent variable without the need for linearity assumptions. Additionally, 0.02, 0.13, and 0.26 values typically represent small, medium, and large effect sizes, respectively [40]. Eta squared assists researchers in better understanding their data and the number of independent variables involved.

According to Table 6, the analyses conducted within the experimental and control groups revealed that both methods were effective. For this purpose, pairwise comparisons (pre-test vs. post-test; post-test vs. retention test; pre-test vs. retention test) were performed using one-way analysis of variance followed by pairwise comparisons (Bonferroni) in both groups, and significant differences were observed within each group. (Consequently, in addition to quantitative analysis, qualitative questions were posed to determine the students' cognitive structures. Such a result is inevitable as new and interesting information is learned more easily. The primary question of interest here is to seek an answer to the "HOW" question, which is to determine how well the students understood the underlying physical processes of the information they learned (HOW IT HAPPENED). Therefore, SEMAT was supported with SEMOEQ.)

Table 7 presents Bonferroni results regarding the differences in mean scores of pre-test and post-test students in the experimental group.

Table 7 Pairwise comparisons of SEMAT achievement scores of experimental group students

Group	Comparison		Mean difference	Standard error	p
Experimental	Pre-test	Post-test	– 1.730	0.143	0.000***
		Retention	– 2.189	0.158	0.000***
	Post-test	Pre-test	1.730	0.143	0.000***
		Retention	– 0.459	0.132	0.003**
	Retention	Pre-test	2.189	0.158	0.000***
		Post-test	0.459	0.132	0.003**
Control	Pre-test	Post-test	– 1.013	0.147	0.000***
		Retention	– 1.462	0.169	0.000***
	Post-test	Pre-test	1.013	0.147	0.000***
		Retention	– 0.449	0.146	0.01*
	Retention	Pre-test	1.462	0.169	0.000***
		Post-test	– 0.449	0.146	0.01*

0.05 ≤ p ≤ 0.10 Marginal significance

*0.01 ≤ p ≤ 0.05 Statistically significant

**0.001 ≤ p ≤ 0.01 High level of statistical significance

***p < 0.001 Very high level of statistical significance [58]

Table 7 shows a very high level of statistically significant difference between the pre-test and post-test scores of the experimental group ($p < 0.001$). Considering the average scores, this difference favors the post-test. Given these results, teaching with the Sun-Earth model developed within the research scope effectively improves student achievement. When the difference between this group's pre-test and retention test scores is examined, a very high statistically significant difference is again determined ($p < 0.001$). Considering the averages, the difference favors the retention test. Additionally, a high level of statistically significant difference is found between the post-test and retention scores of the students ($0.001 \leq p \leq 0.01$). Considering the averages, the difference favors retention. When all findings are considered, it can be said that the teaching with the Sun-Earth model developed within the scope of the research serves its purpose.

In the control group, while a very high level of statistically significant difference is found between the pre-test and post-test scores ($p < 0.001$), considering the averages, this difference favors the post-test. When the difference between this group's pre-test and retention scores is examined, a very high level of statistically significant difference is also found ($p < 0.001$). Considering the averages, the difference favors the retention test. Additionally, a statistically significant difference is found between the post-test and retention scores of the students ($0.01 \leq p \leq 0.05$). Considering the averages, the difference favors retention. When all findings are considered, it can be said that teaching with the Sun-Earth model developed within the scope of the research serves its purpose.

The achievements of the experimental and control groups in the applied tests were examined within each group. Following this examination, comparisons were made between the groups and the measurements to determine (if any) the differentiation between the teaching method with the physical models developed within the scope of the research and the teaching methods used in the National Education curriculum. For this purpose, a two-way analysis of variance technique for mixed measurements was used to compare the measurements and the groups.

We examined the findings related to the change in the groups' SEMAT scores, and Table 8 shows the results of the two-way analysis of variance.

According to Table 8, the results obtained from the achievement tests applied to the groups show a significant difference between the groups without separating the achievement scores into pre-test, post-test, and retention test ($F = 7.625$, $p < 0.05$). Additionally, this difference is also determined separately for each of the three tests when evaluating achievement between the groups ($F = 168.445$, $p < 0.05$). These results indicate that the teaching methods and techniques used in both groups affect achievement. A post-hoc (Scheffe) multiple comparison analysis was conducted to determine which specific pairs of subgroup mean scores contribute to this difference. The results of this analysis are presented in Table 9.

When analyzing the data, a significant result favoring the experimental group is observed across all three applications. For a more precise interpretation of this change, refer to Fig. 9.

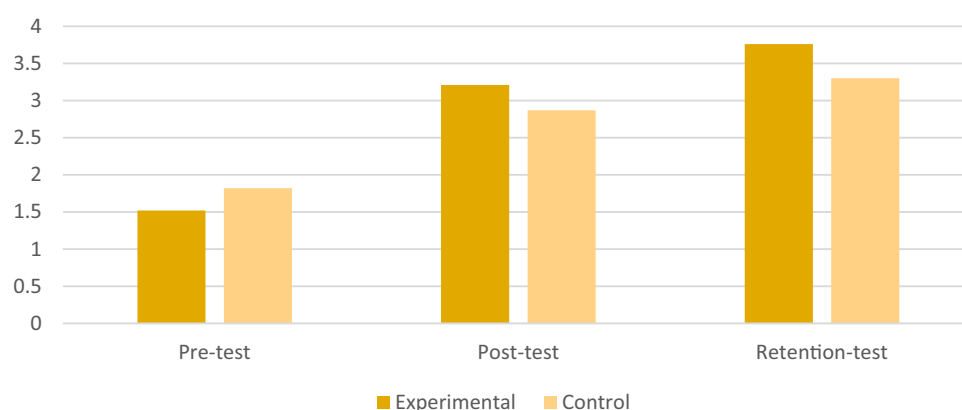
Upon examining Fig. 9, an increase is observed in both groups in the comparisons between the pre-test and post-test and between the post-test and retention test. However, upon closer inspection, it is evident that although the experimental group had lower scores than the control group in the pre-test, they exhibited a higher increase in the post-test

Table 8 ANOVA results for SEMAT achievement scores of experimental and control group students

Source of variance	SS	df	MS	F	p	η^2
Between Groups	321.298	145				
Group (Experimental-Control)	15.642	1	15.642	7.625	0.006	0.214
Error	305.656	149	2.051			
Within Groups	458.193	300				
Measurement (Pre-Post-Retention)	228.941	2	114.470	168.445	0.000	0.337
Group*Measurement	27.342	2	13.671	20.117	0.000	0.450
Error	202.512	298	0.680			
Total	779.491					

Table 9 Post-Hoc (Scheffe) results for experimental and control groups

Measurement	Group	N	1	2	3
Pre-test	Experimental	74	1.54		
	Control	78	1.82		
Pos-test	Experimental	79		3.21	
	Control	76		2.86	
Retention test	Experimental	79			3.76
	Control	77			3.03

Fig. 9 Graph of changes based on group and measurement

than the control group. Similarly, there was a more significant increase in the retention test in the experimental group compared to the control group. It can be inferred that the difference between the two groups arises from these instances.

In conclusion, in teaching any concept or phenomenon, it is natural to achieve a certain degree of learning, regardless of the method or technique employed. This is unavoidable. However, the critical aspect is determining the extent and depth of the effectiveness of the employed method. Examining Fig. 9, it is evident from the SEMAT results that learning has occurred (naturally) in both the experimental and control groups. Although the changes appear minor according to the graph, slight increases or decreases are significant when considering a 5-point test. According to Fig. 9, the control group started with a score of 1.82 before the teaching process, while the experimental group began with a score of 1.52. The study commenced with a 6% advantage in favor of the control group. However, the final test results showed an advantage for the experimental group, which scored 3.21 compared to the control group's 2.87. This resulted in a 6.8% advantage for the experimental group post-instruction. Compared to the pre-test scores, there was a 12.8% improvement in favor of the experimental group.

After 8 weeks, the retention tests revealed a further increase in this difference. It can be interpreted that the experimental group internalized the knowledge and developed mental models from the instruction provided through the study's developed model. The experimental group reached a score of 3.76, while the control group achieved a score of 3.30. The retention rate in the experimental group was 75.2%, whereas in the control group it was 66%. Considering the change from pre-test to retention, the experimental group showed a 44.8% increase, while the control group exhibited a 29.6%

increase. As previously mentioned, some degree of increase is inevitable regardless of how the teaching is conducted. The crucial point is the magnitude of this increase. Additionally, open-ended questions were posed to the students to determine whether these improvements in both groups were superficial or deep. Understanding the formation of seasons likely involves knowledge of axial tilt, the Earth's spherical shape, and, most importantly, the Earth's orbital (revolution) movement around the Sun. However, learning how these factors physically influence the formation of seasons requires a deeper understanding. This study was designed with this perspective in mind.

5.2 Findings of descriptive analysis of students' responses to open-ended questions (SEM0EQ)

To determine whether deep learning occurred, we asked the students "HOW." We also informed them that they could use drawings while providing their explanations. We compared the data obtained from SEMAT with the SEM0EQ data. During the content analysis, we established categories. These categories and correct responses are given in Table 10.

When SEM0EQ data is analyzed, both groups are expected to give a high percentage of correct answers in categories 2 and 3. This is a predictable result, as the effects of these factors on the formation of seasons are emphasized at every stage of education throughout the teaching process. The distinguishing categories here are 1, 4, and 5. The variations in these categories can be interpreted as students conceptualizing the underlying physical phenomena responsible for the formation of seasons. Category 7, on the other hand, represents a misconception. In the control group, some students believed that the Earth's distance from the Sun affects the seasons, even after the instructional process had concluded.

Upon examining the table, it becomes evident that the explanations provided by students in the experimental group, specifically in categories 1, 4, and 5—demonstrating their understanding of the physical phenomena based on scientific principles—were learned more effectively and retained better than those in the control group. When these findings are interpreted in conjunction with the SEMAT, it can be concluded that an effective teaching method was applied to the experimental group.

The scope of the open-ended question prepared for students is about how the seasons are formed. To ensure objectivity in the study, responses given by ten students systematically selected from each group were meticulously analyzed for the pre-test (P), post-test (S), and retention test (R). Attention was paid to systematic selection by skipping every eighth student, starting from the first student selected. The responses provided by the students are presented in Table 11.

Table 10 shows that the pre-test scores for categories 6, 7, and 8 are significantly high in the experimental and control groups. This indicates that both groups possess conceptual misunderstandings regarding the formation of seasons. Data obtained during the study, particularly in categories 6 and 7, reveal a misconception about the elliptical movement of the Earth around the Sun and the effect of the hemisphere facing the Sun on the formation of seasons. However, the distance between the Earth and the Sun does not influence the formation of seasons. This is one of the most significant misconceptions in these categories. When examining the post-test results, a decline is observed in both groups, but it was found that all members of the experimental group eliminated these conceptual misunderstandings. The retention test administered to the experimental group showed that the models used prevented the reoccurrence of these conceptual misunderstandings. In contrast, in the control group, the rate of conceptual misunderstanding increased from 3.4% to 6.9%, indicating that the teaching method used in the control group resulted in relatively low retention of learned knowledge. These findings suggest that adequate, appropriate models that do not cause conceptual misunderstandings enhance knowledge retention in education.

Table 10 reveals that Item 8 contains responses not included in the first seven items and is scientifically inaccurate. For instance, responses suggesting that the Sun revolves around the Earth or that the Moon influences the formation of seasons due to its orbit around the Earth are considered scientifically unfounded and have therefore been categorized under "Other." In this category, several conceptual misunderstandings were observed in both groups during the pre-test. However, in the post-test, the experimental group's misunderstanding rate decreased to 3.4%, while the control group's rate dropped to 6.9%. The decline in the experimental group was comparatively more pronounced. Nevertheless, contrary to the study's objectives, the retention test increased the experimental group's rate to 6.9%.

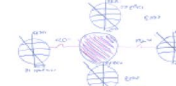
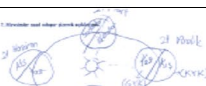
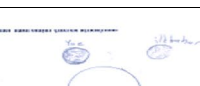
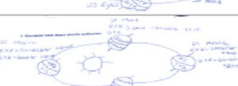
Table 11 indicates that the students were systematically selected to avoid random errors. Upon careful examination of the data, it is observed that the readiness of the groups before the teaching process was equal (The more detailed version of the table is provided in the "Appendix" section for further examination.). The most significant change observed in the experimental group after the teaching process was the inclination of axes in their drawings. Upon closer inspection of the drawings, it was determined that the drawings of the experimental group were scientifically more accurate. The inclination of the axes, not observed in the drawings of the pre-test, was reflected in the post-test and retention tests.

Table 10 Pre-post and retention test percentage distributions on the topics How are seasons formed?

No	Categories	Explanation of categories	Groups	Pre-test %	Post-test %	Retention-test %
1	Angle difference of rays ^a	Explanations stating that the angle formed between the Sun's rays, which are parallel to the Earth, and the Earth's surface changes, resulting in a variation in the amount of energy falling per unit area	Experimental Control	37.9 30.6	91.8 40.8	85 34
2	Axial tilt ^a	Explanations state that the Earth has an axial tilt of 23.5°	Experimental Control	62.1 54.4	100 91.8	96.5 71.4
3	Earth's revolution ^a	Explanations stating that the Earth revolves around the Sun	Experimental Control	79.3 74.8	100 95.2	89.6 84.6
4	Earth's shape ^a	Explanations stating that the Earth has a spherical shape	Experimental Control	3.4 0	62.1 23.8	44.8 17
5	Amount of energy per unit surface ^a	Changes in energy levels per unit area in the region where the Sun's rays fall on the Earth's surface, based on Categories 1, 2, 3, and 4	Experimental Control	0 0	91.8 13.6	81.6 3.4
6	Elliptical movement	Explanations stating that the Earth follows an elliptical orbit around the Sun	Experimental Control	10.3 23.8	0 3.4	0 6.9
7	Earth's position	Responses indicate the influence of the Earth's distance from the Sun and the Earth's hemisphere facing the Sun	Experimental Control	6.9 13.6	0 3.4	0 10.2
8	Other	Responses that are not included in the first seven categories and are nonsensical	Experimental Control	17.2 20.4	3.4 6.9	6.9 6.9
9	Not answer	Unanswered responses	Experimental Control	0 0	0 0	3.4 0

^aScientifically accurate

Table 11 Student explanations regarding “how are seasons formed?”

Stu	App	Experimental gr.	Control gr.
10	O		YANITSIZ
	S		
	K		
26	Ö	YANITSIZ	
	S		
	K		YANITSIZ
34	O		
	S		
	K		
58	Ö		
	S		
	K		

*Ö: Pre-test; S: Post-test; K: Retention-test; YANITSIZ: UNANSWERED; Aralık: Dec; Eylül: Sep; Mart: March; Haziran: June; Güneş: Sun; Dünya: Earth; İlkbahar: Spring; Sonbahar: Autumn/Fall; Kış: Winter; Yaz: Summer; KYK: Northern Hemisphere-NH; GYK: Southern Hemisphere-SH

Responses from selected students, specifically those numbered 10, 26, 34, and 58 from the experimental and control groups, were analyzed.

For student 10, it was observed that the experimental group student considered the effect of Earth's 23.5-degree axial tilt on the seasons. In contrast, the control group student explained the seasons without considering this tilt.

For student 26, the experimental group emphasized axial tilt, while the control group students provided a rote explanation without considering axial tilt. Additionally, when examining the retention test, the control group student failed to explain, whereas the experimental group student successfully answered by referencing axial tilt.

For student 34, the experimental group student's response included a more detailed representation of axial tilt. Furthermore, the retention test results demonstrated that the knowledge gained by the experimental group student was retained, indicating the effectiveness of the intervention.

Student 58's responses from the experimental group showed some prior knowledge before the study. However, this student provided more organized and meaningful explanations following the survey. In contrast, the control group student's responses showed no significant improvement between the pre-and post-intervention phases.

The results obtained from the two separate teaching processes for the experimental and control groups, based on their achievements in the study, are as follows:

The descriptive statistics of the experimental and control groups appear to be similar. Additionally, it was found that the classes' readiness before the teaching process was low.

In the control group, although students referring to the inclination of axes were identified, their drawings indicated that they could not think in three dimensions, only at the level of learning from two-dimensional sources. Additionally, in the experimental group, some students' explanations of the formation of seasons, mentioning "sphericity" and "orbital motion," indicate that they were able to reach levels of understanding, application, and even analysis beyond the level of recalling the formation of seasons, as they were able to reflect this in their drawings. Unlike multiple-choice test analyses, open-ended questions demonstrate that the explanations and drawings made by students in the experimental group on the formation of seasons were more straightforward compared to those of the control group. Moreover, in both groups, some students clearly did not understand the formation of seasons, but this number was higher in the control group.

6 Discussion and conclusion

6.1 Discussion

The findings of this study are consistent with the existing literature that highlights the importance of materials and models in education and are supported by recent studies. The impact of model use on student achievement and its role in teaching scientific concepts are explained in our research. The nature of the seasons is considered one of the most challenging scientific concepts to be grasped. A recent study found that students often struggle to understand the scientific basis of this topic, mainly when taught using both augmented reality and physical models. The literature also reports similar findings [43–47].

Our research shows that the CAEFUS model helped students better understand the relationship between axial tilt and the Earth's orbital motion. This result aligns with current literature, confirming that using models significantly enhances students' conceptual understanding. Recent research also supports the effectiveness of model use. Chen and Lin noted that using models to understand scientific concepts contributed to students' conceptual change and improved their learning processes [48]. Similarly, Soliman and Moorthy reported that models enhanced students' performance in understanding physical phenomena [49]. These findings reinforce the educational impact of model use.

Students' difficulties in understanding the relationship between axial tilt, the Earth's orbit, and the amount of energy affecting seasons are also supported by current research. Sánchez et al. addressed the challenges students face in understanding concepts related to the formation of seasons and suggested that these difficulties could be overcome with appropriate teaching strategies [50]. The model developed in this research and the study is aligned with these findings. The students' struggles to understand axial tilt and energy about the formation of seasons were examined by Zhang and Liu and it was shown that effective models play a crucial role in addressing these issues. [51].

Recent literature examining the impact of models on learning reveals that model-based learning strategies effectively enhance students' conceptual understanding. Karataş and Yıldız noted that model-based learning contributes to a deeper understanding of knowledge [52]. Nair et al. also emphasized that models are a crucial tool for helping students grasp physical concepts [53]. Understanding physical phenomena is also supported by current research. Patel and Sharma highlighted the contributions of models in helping students understand and analyze physical events, demonstrating

that model use is effective in understanding the mechanisms behind scientific concepts [54]. Therefore, the findings of this research support the existing literature and serve as a foundation for future studies.

These findings support the study's results and reinforce the importance of models in education. Additionally, they can serve as a basis for future research and provide new approaches for improving educational materials and methods.

7 Conclusion

In the new curriculum introduced by the Ministry of National Education, models are recommended as one of the most effective methods for teaching astronomy concepts. However, the models must closely align with reality, crucial in shaping students' mental schemas. In this context, the CAEFUS model, developed within the scope of this research, was employed to evaluate its impact on student achievement. It was observed that the CAEFUS model significantly enhanced students' understanding of the formation of seasons. The model helped students comprehend the role of axial tilt, the Earth's orbital motion around the Sun, and the effect of the amount of energy on the formation of seasons. The visual and concrete representations provided by the models enabled students to learn these concepts more effectively and take a more active role in the learning process.

This research demonstrated that students quickly learned that the Earth's distance from the Sun does not influence the formation of seasons through their explanations and visual representations. Similarly, Galano emphasized that the Earth's distance from the Sun does not affect the seasons [55]. Trumper found that while students from elementary to university levels acknowledged that the Earth's axial tilt causes seasonal changes, they struggled to explain the temperature differences between summer and winter [56]. Henriques also found that while students understood that the Earth's distance from the Sun did not affect the seasons, they had difficulty explaining why [57]. However, in this study, students in the experimental group better understood the relationship between the axial tilt and the Earth's orbital motion with the nature of the seasons and were able to depict it visually. This suggests that the CAEFUS model is effective in student learning. Galano et al. noted that while students could understand the effects of axial tilt and the Earth's orbit on the nature of the seasons, they struggled to connect the energy received by the Earth with the varying angles of sunlight [55]. Thus, using appropriate models is critical in deep scientific learning instead of superficial understanding. Consequently, the CAEFUS model effectively influences students' explanations and visual representations of the nature of seasons.

Additionally, an increase in achievement was observed in both the experimental group, where models were used, and the control group, where only textbook images and text were utilized. However, this increase was more pronounced in the experimental group. These findings indicate that understanding the nature of seasons—particularly the concept of a tilted axis and the Earth's orbit around the Sun—is one of the most challenging concepts to grasp. Students in the experimental group developed different perspectives when interpreting the nature of seasons compared to the control group, as the model presented to them was as free of conceptual misconceptions as possible. Understanding the role of the amount of energy on the nature of seasons helped students better grasp the concepts of Earth's axial tilt, orbit, and even its sphericity. As a result, students' explanations and visual representations of the nature of seasons helped clarify the relationship between the transfer of solar energy and the Earth's surface area, thereby enhancing student achievement.

The simultaneous occurrence of many physical events shapes the nature of the seasons. For example, questions such as how seasons would be affected if the Earth were not spherical, if there were no axial tilt, or if the Earth did not orbit the Sun were explored using models. Exploring these questions helped students establish relationships between hypothetical situations and current conditions. In this context, the models used helped measure and observe different conditions, improving students' abilities to analyze and imagine, thus enabling them to explain the nature of seasons more meaningfully and represent it visually.

Moreover, the model developed in this study emphasizes the interplay of scientific and physical principles essential for teaching the nature of seasons. While specifically designed for astronomy education and explaining the concept of seasons, the model's design elements—such as representing energy distribution on surfaces and employing scaled physical models for teaching—have the potential to be adapted to other scientific fields. The model can be adapted to various fields, such as geography and climate studies, to explain solar energy distribution across different latitudes or seasonal variations in climate zones. In physics education, the principles of light incidence and energy transfer used in the model can serve as a teaching tool in optics or thermodynamics courses. Additionally, in environmental science, concepts of energy balance and solar radiation can be demonstrated to explain climate change and its effects on ecosystems.

In summary, with necessary adaptations, the CAEFUS-based model can be utilized in various educational contexts to facilitate the understanding of complex scientific concepts beyond the study of seasons.

8 Recommendations

Educational materials and models should be continually updated and supported by innovative approaches to be more effective. Future research could benefit from comparing the effects of different types of models and improving model design. Additionally, various methods and tools should be employed when teaching students scientific concepts. For example, teaching methods supported by new technologies such as augmented reality and simulations should be explored.

Longer-term studies should be conducted to assess the long-term effects of models and materials. Such research could help measure how long students retain the concepts they learn and how persistent their understanding is. Educational programs should provide students with superficial information and deep content that helps them understand the physical mechanisms behind this information. This will help improve students' conceptual understanding and enhance their ability to apply their knowledge.

These recommendations could guide future research and practices aimed at increasing the effectiveness of educational models and supporting student achievement.

Author contributions M.G.S.: literature review, determining the methodology, data collection, data analysis, report writing, writing the discussion, validity and reliability studies. H.K.: Research design, model development, consulting.

Funding The authors have no conflicts of interest to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Data availability Data is provided within the manuscript or supplementary information files but if there is a request for all the obtained data, they can be provided by the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate The protocol was approved by the Provincial Directorate of National Education of the Ministry of National Education of Turkey in accordance with the "Ministry of National Education Directive on Permission and Implementation of Research and Research Support to be Conducted in Schools and Institutions" and the "approval of the relevant directorate dated 24.05.2019 and numbered 10294247. Before the commencement of this study, formal approval was obtained from the Ministry of National Education of Turkey, the authorized body in the country. Following this, consent was obtained from the families through a consent form for the implementation of the measurement. Participants were informed that no information that could identify the school or the students would be disclosed during the data analysis process. Signed consent forms were collected from all participants.

Consent for publication The need for informed consent was waived by the Directorate of National Education.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Appendix

Figures 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and Table 12

Fig. 10 Differences in the distribution of sun rays on the earth's surface with flat plate and spherical surface

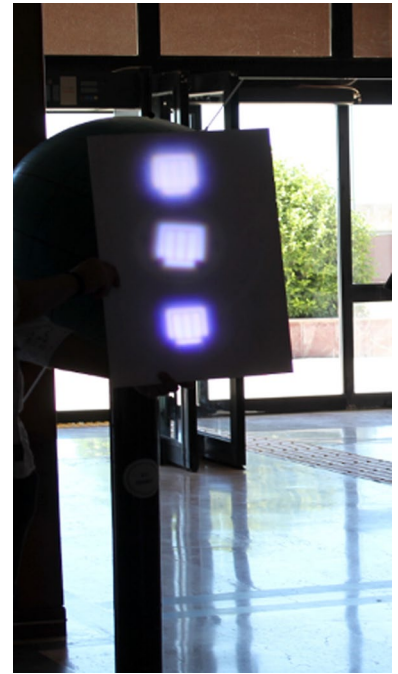


Fig. 11 Calculate the energy falling on a unit surface as the Earth orbits the Sun and reaches different regions at different angles



Example Lesson Plan

Topic: Teaching the Formation of Seasons

Target Audience: 8th Grade Students

Lesson Duration: Planned for 2 class periods

1st lesson plan: formation of seasons—basic concepts and model setup

Objective: Students will learn the key factors affecting the formation of seasons and the role of energy falling per unit area. They will visually explore the effect of these factors on the seasons through a model.

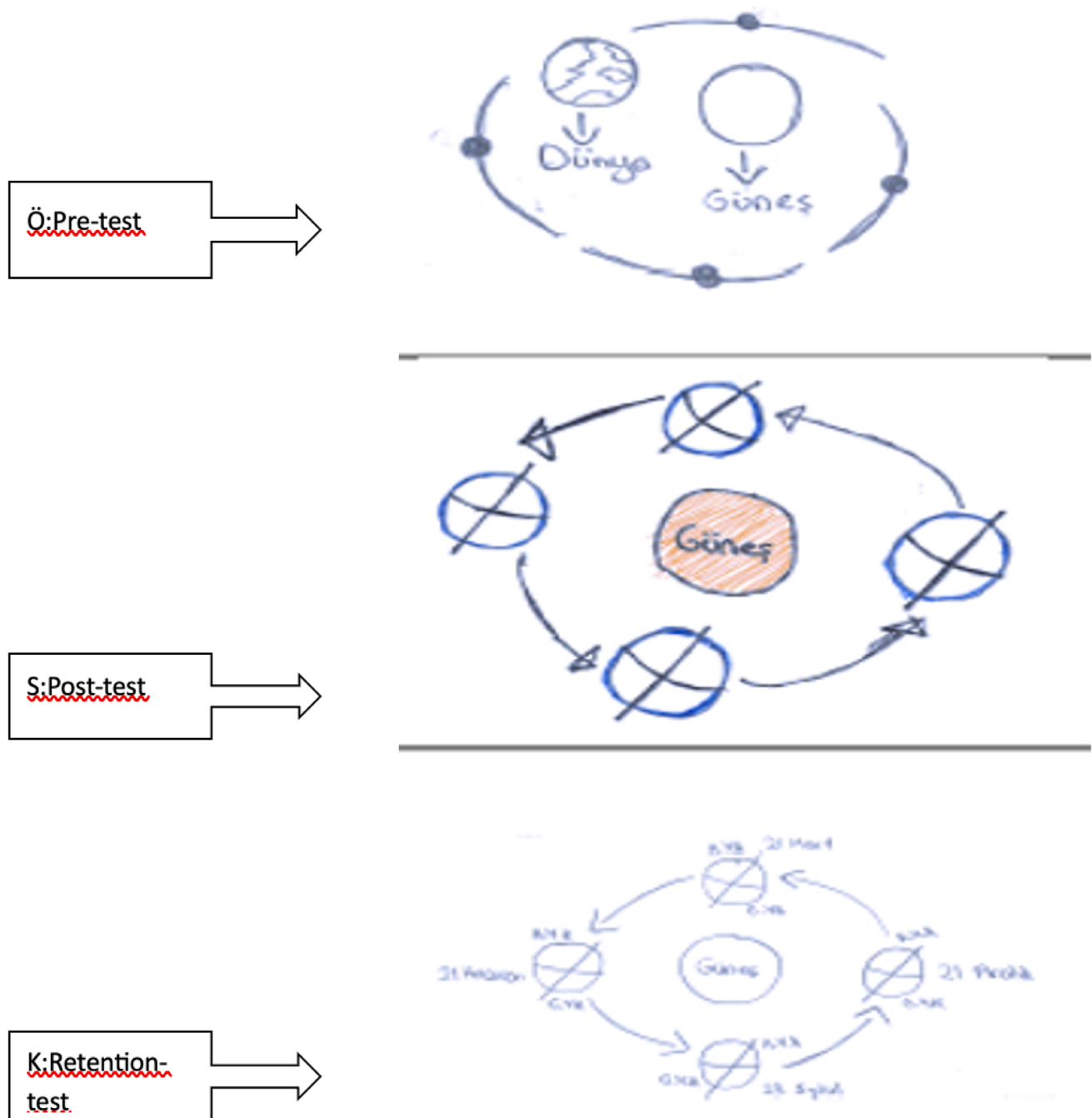


Fig. 12 Experimental group student number 10

- a. Introduction (5 min)
- a Brief Introduction: A short conversation with students about why seasons change.

Introduction of Objectives: Students are expected to learn about the factors influencing the formation of seasons and how these factors relate to energy falling per unit area.

- b. Teaching New Information (10 min).
- b Explanation of the Formation of Seasons:
- c The effects of the Earth's shape, axial tilt, and movement around the Sun on seasons will be explained.

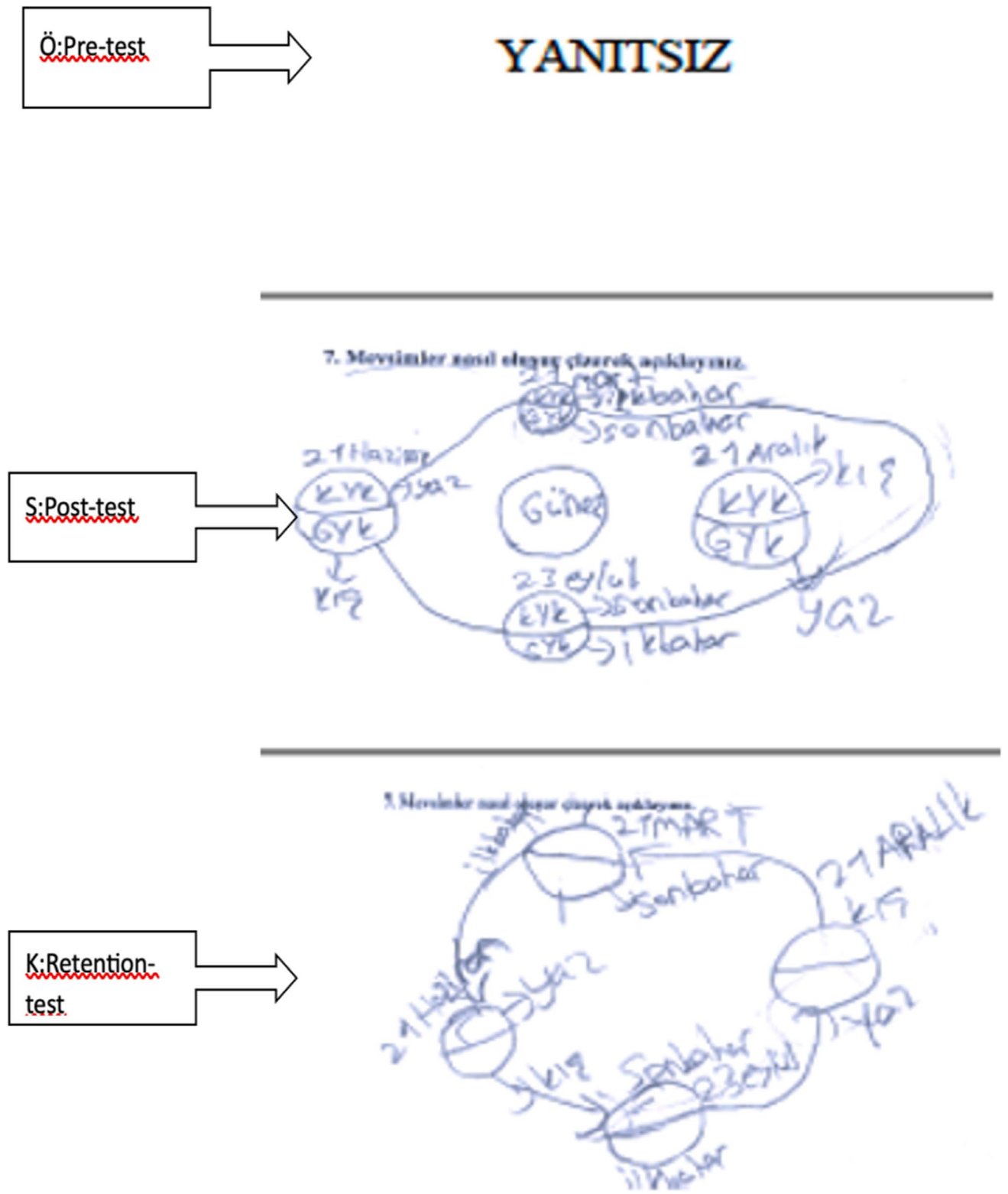


Fig. 13 Control group student number 10

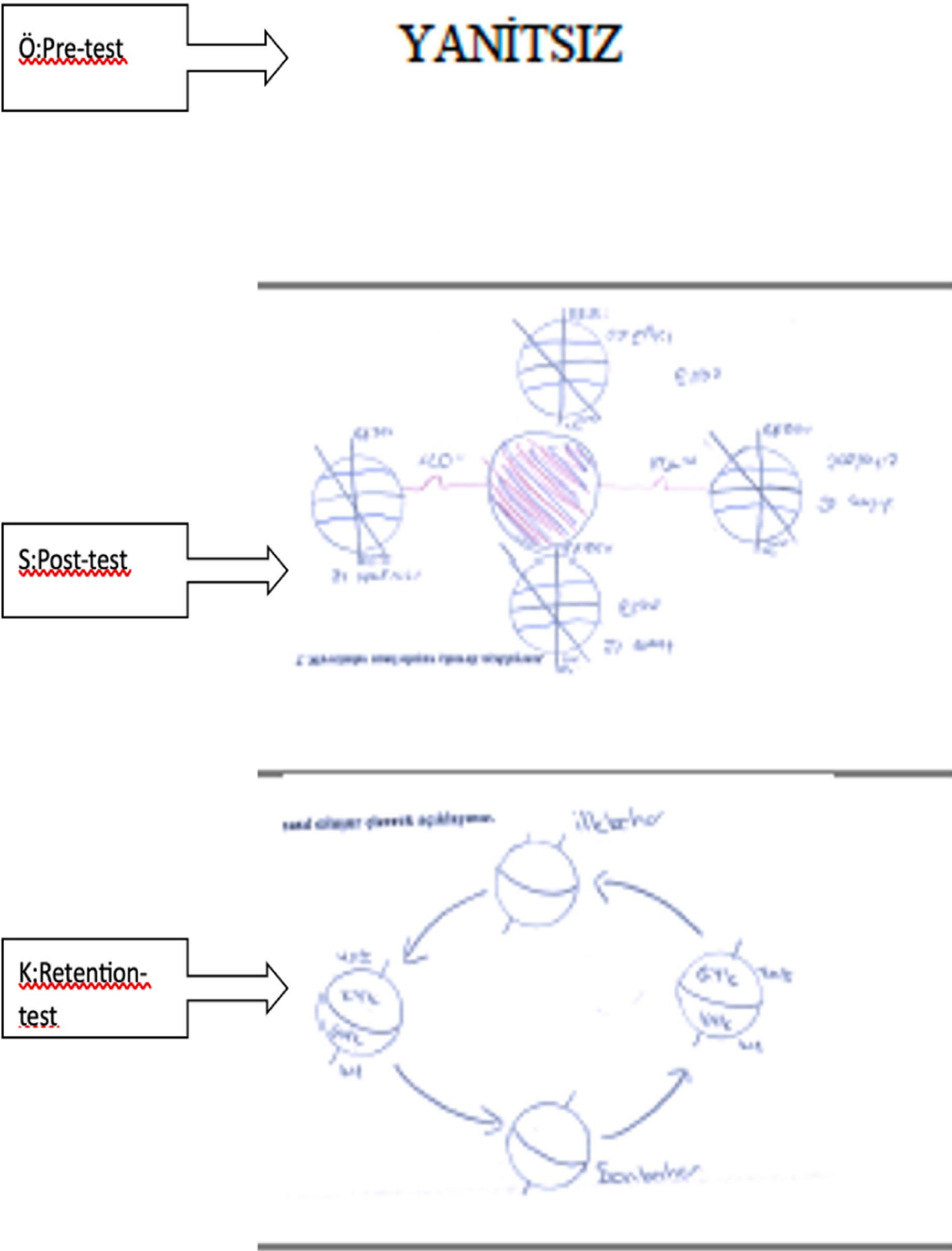


Fig. 14 Experimental group student number 26

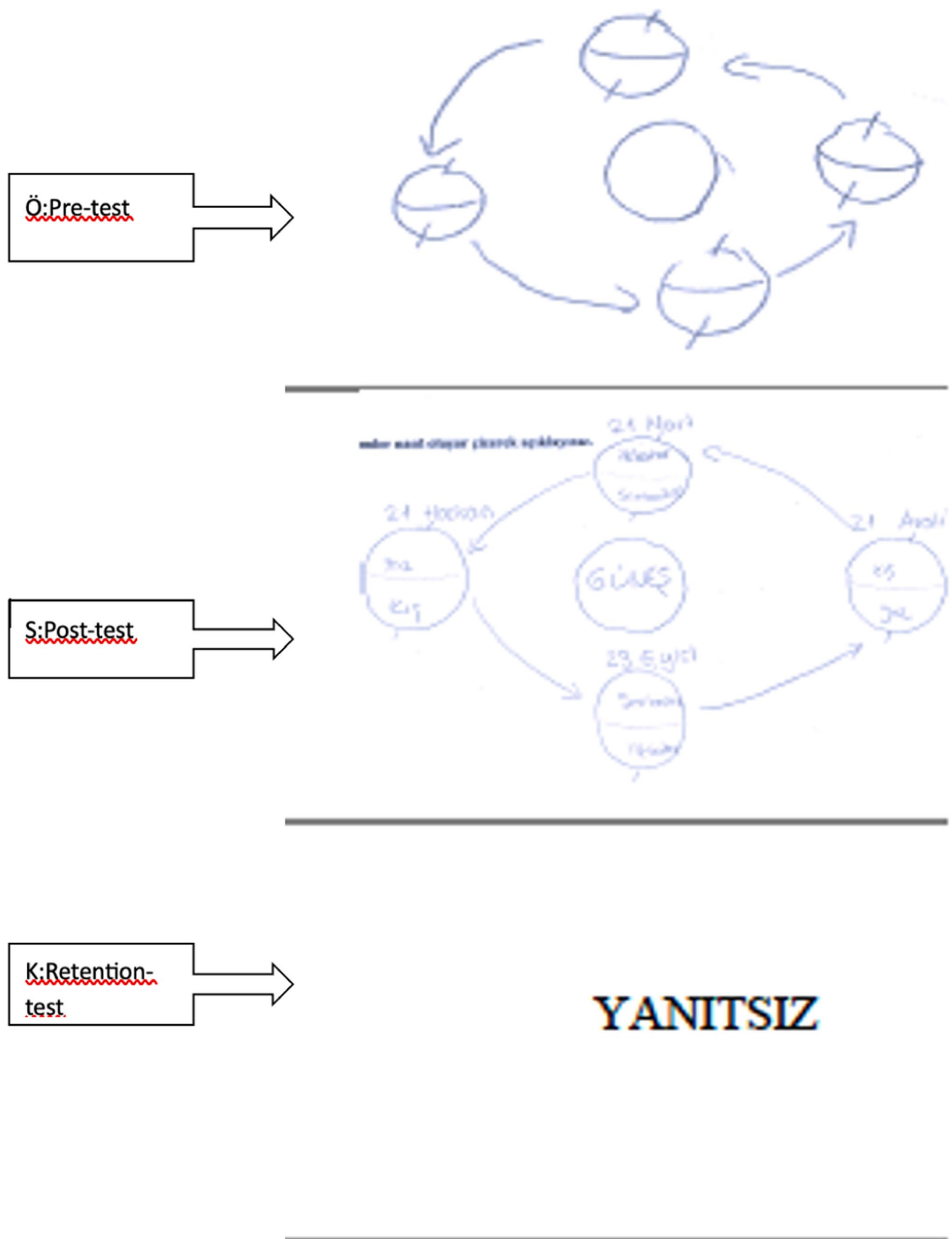


Fig. 15 Control group student number 26

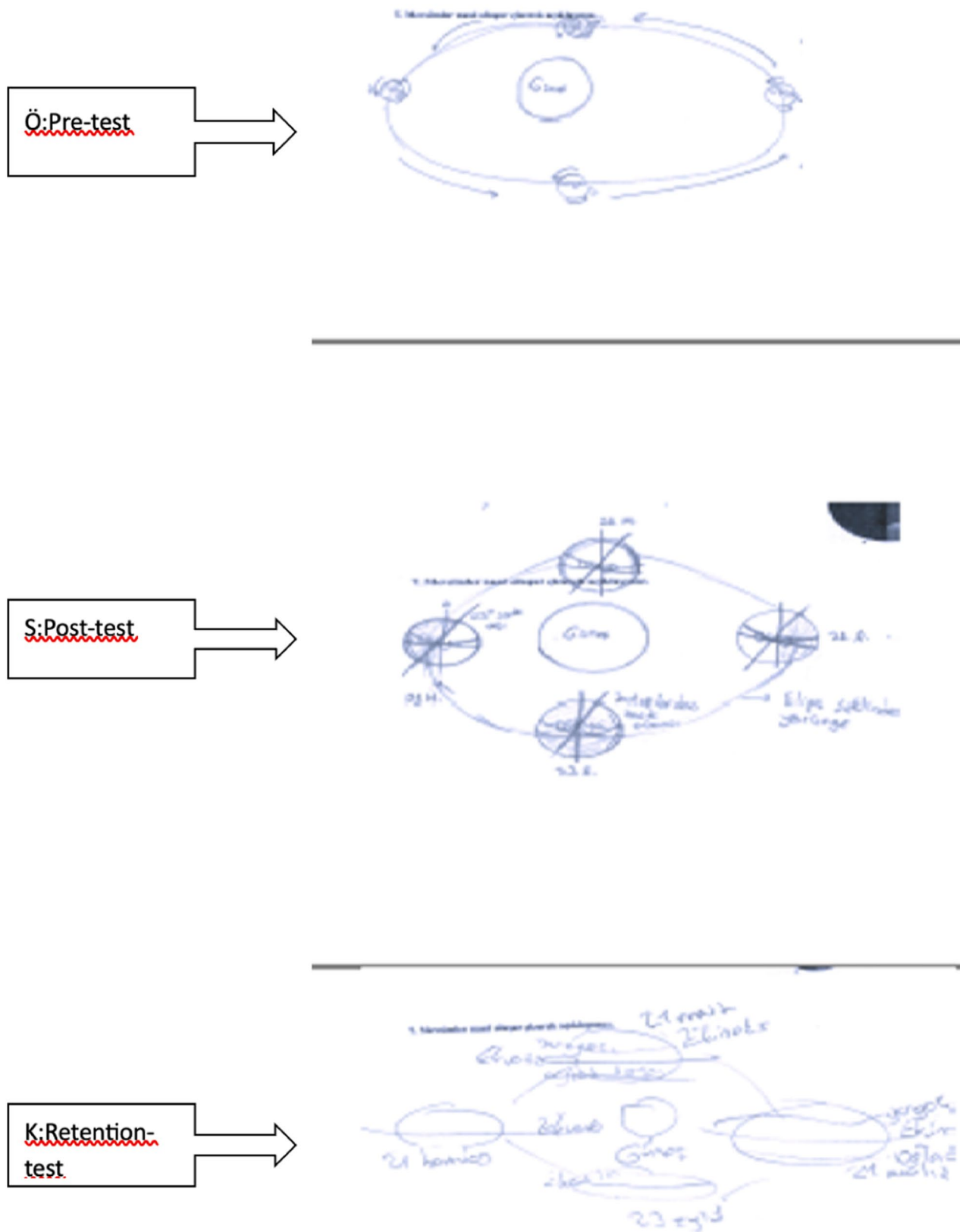


Fig. 16 Experimental group student number 34

Fig. 17 Control group student number 34

Ö:Pre-test



S:Post-test



K:Retention-test



- d The relationship between the amount of energy falling per unit area and the formation of seasons will be clarified.
- c. Model Setup (15 min).
- e Plates and Flashlight System:
- f Four plates representing the equinoxes and solstices will be placed.
- g Three strong flashlights (simulating the Sun) will be set up.
- h Students will observe different Earth positions around the model and understand how rays of light fall on various regions.

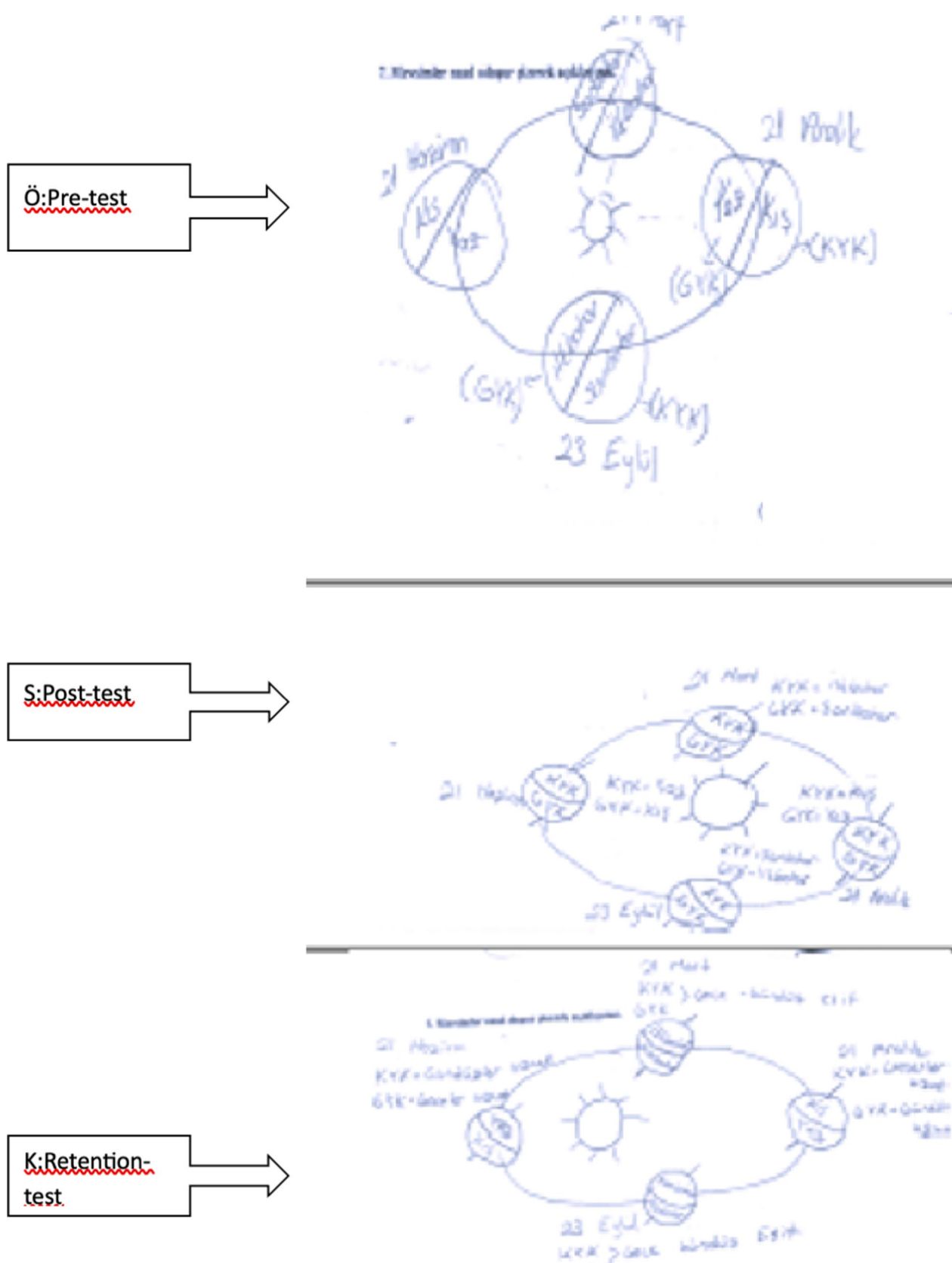


Fig. 18 Experimental group student number 58

Fig. 19 Control group student number 58

Ö:Pre-test →



S:Post-test →



K:Retention-test →

**Table 12** Translate the Turkish terms into their English equivalents

Turkish	English	Turkish	English	Turkish	English
Öntest (Ö)	Pre-test	Aralık	December	İlkbahar	Spring
Sontest (S)	Post-test	Eylül	September	Yaz	Summer
Kalıcılık testi (K)	Retention-test	Mart	March	Sonbahar	Autumn/Fall
YANITSIZ	UNANSWERED	Haziran	June	Kış	Winter
Güneş	Sun	Kuzey Yarım Küre (KYK)	Northern Hemisphere (NH)	Güney Yarım Küre (GYK)	Southern Hemisphere (SH)
Dünya	Earth				

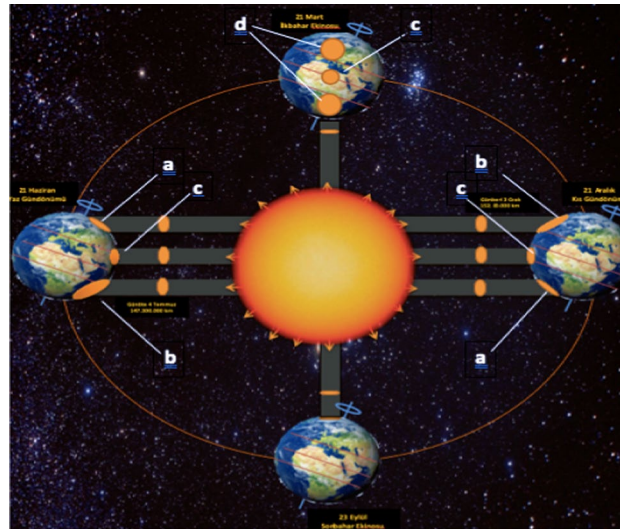


Figure A The Earth's orbit around the Sun is one of the key factors in the formation of seasons. When combined with the Earth's axial tilt, it results in different seasons occurring in various regions of the world at different times of the year. Imagine if the Earth had no axial tilt and simply orbited the Sun, rotating around its axis as it does now. For 364 days and 6 h, each region would experience the same season. The equator would remain hot, the poles would be extremely cold, and the temperate regions would experience spring year-round. Alternatively, if the Earth had an axial tilt but did not orbit the Sun, one hemisphere would experience summer for 365 days and 6 h, while the other hemisphere would experience winter. In both cases, the entire year would consist of a single season

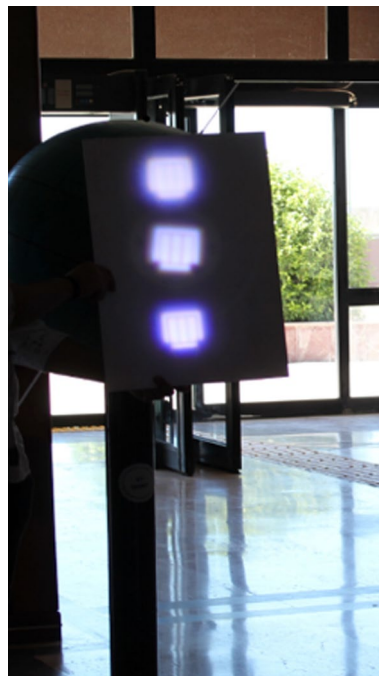


Figure B One of the factors that affects the distribution of solar radiation on the Earth's surface is the Earth's spherical shape. If the Earth had a flat surface, the same season would occur across the entire surface because the same amount of energy would fall on each unit area. However, as shown in Fig. C, because the Earth has a spherical shape, the solar rays fall more directly on the equator, causing a smaller area to be exposed to a higher amount of energy, leading to a different level of heating. In contrast, in the northern and southern hemispheres, where the rays come

at a wider or narrower angle (greater than 90 degrees), the energy spreads over a larger surface area, reducing the amount of energy per unit area. Therefore, the amount of energy falling per unit area is a critical factor in the formation of seasons. This is why different seasons occur in different regions of the Earth. For instance, the poles experience constant cold seasons, the equator experiences hot seasons, and temperate regions experience four distinct seasons

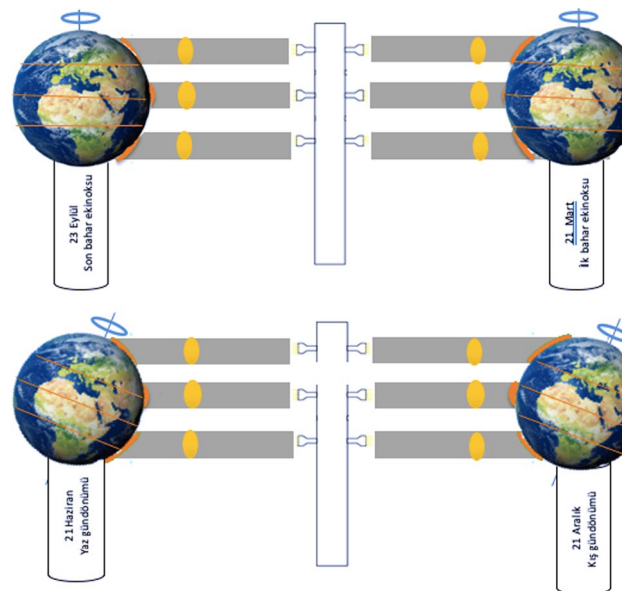


Figure C The angle between the solar rays and the Earth's surface on different dates is shown, illustrating the area over which energy is distributed. Due to the Earth's movement around the Sun, along with its axial tilt and spherical shape, the amount of energy falling on different regions of the Earth's surface changes throughout the year, resulting in the formation of seasons

Three flashlights representing the Sun are placed here.



Figure D Representation of the Sun and the Earth's position during the equinoxes. In all positions, the Earth's axis is aligned with the North Star. The Earth's axes point towards the North. The distribution of solar radiation reaching the equator and the poles in all positions is examined and interpreted

d. Student Participation (5 min).

Question and Answer: Based on the model, questions such as “What factors influence the change of seasons?” will be asked to help students reinforce their understanding.

e. Closing and Homework (5 min).

e Assigning Homework: Students will be asked to create a drawing or a small model related to the information they learned today. This model should visually represent the Earth’s movement around the Sun and the formation of seasons.

2nd lesson plan: formation of seasons—application of the model and student work

Objective: Through their creative work, students will reinforce the information they learned in the first lesson and better understand the formation of seasons.

a. Introduction (5 min)

a Review of the Previous Lesson: Students will recall the concepts they learned in the previous lesson.

b Introduction of the New Activity: Today, students will visually express their learned information.

b. Students’ Work (25 min)

c Making a Model or Drawing:

d Students will create a model or drawing that demonstrates the formation of seasons using the materials provided or their creative ideas.

e Students should show how the Earth’s shape, axial tilt, and movement around the Sun affect the formation of seasons in their models.

c. Sharing Student Work (10 min)

f Presentation and Discussion:

g Students will share their models or drawings with the class.

h Each student will explain how they represented the formation of seasons in their work.

i Students will review and give feedback on each other’s work.

d. Closing (5 min)

j Summarizing What Students Learned:

k We will review the essential factors that influence the formation of seasons with the students.

l We will discuss how the models used in the lesson helped us understand the concepts.

References

1. Genovese JEC. Piaget, pedagogy, and evolutionary psychology. *Evol Psychol.* 2003. <https://doi.org/10.1177/147470490300100109>.
2. Bransford JD, Schwartz DL. Rethinking transfer: a simple proposal with multiple implications. *Rev Res Educ.* 1999;24(1):61–100. <https://doi.org/10.3102/0091732X024001>.
3. Langmuir CH, Broecker W. How to build a habitable planet: the story of earth from the big bang to humankind. Beijing: Princeton University Press; 2012. <https://doi.org/10.2307/j.ctt1r2f84>.
4. Cho HS, Nam J. Analysis of trends of model and modeling-related research in science education in Korea. *J Korean Assoc Sci Educ.* 2017;37(4):539–52.
5. De Paor DG, Drodevic MM, Karabinos P, Burgin S, Coba F, Whitmeyer SJ. Exploring the reasons for the seasons using google earth, 3D models, and plots. *Int J Digit Earth.* 2017;10(6):582–603.
6. Kim HJ. A comparative study of the Bohr atomic model and the spectrum of atomic hydrogen in chemistry and physics curricula. *Korean Soc Sch Sci.* 2015;9(2):94–100.
7. Plummer JD, Maynard L. Building a learning progression for celestial motion: exploring students’ reasoning about the seasons. *J Res Sci Teach.* 2015;51(7):902–29.
8. Hsu Y. Learning about seasons in a technologically enhanced environment: the impact of teacher-guided and student-centered instructional approaches on the process of students’ conceptual change. *Sci Educ.* 2008;92(2):320–44.
9. Starakis I, Galani A, Angeliki L. The use of “Scratch” in geography: the teaching of seasons. In: 9th Panhellenic Conference of ICT Educators. 2017. pp. 128–36.

10. Van Loon MH, Dunlosky J, Van Gog T, Van Merriënboer JJ, De Bruin AB. Refutations in science texts lead to hypercorrection of misconceptions held with high confidence. *Contemp Educ Psychol*. 2015;42:39–48.
11. Kolchin IS, Miroshnichenko AS, Kadeeva OE, Syritsyna VN. 3D modeling as a tool for gamification of the process of studying science disciplines. *Sovremennyye Problemy Nauki i Obrazovaniya*. 2022;6(1):45. <https://doi.org/10.17513/spno.32256>.
12. Wilson MD, Boag RJ, Strickland L. All models are wrong; some are useful, but are they reproducible? *Comput Brain Behav*. 2019;2:239–324. <https://doi.org/10.1007/s42113-019-00054-x>.
13. Trumper R. Factors affecting students' junior high school students' interest in physics. *J Sci Educ Technol*. 2006;15(1):47–58. <https://doi.org/10.1007/s10956-006-0355-6>.
14. Katarina S, Pavlin J. Improvements in teachers' knowledge and understanding of basic astronomy concepts through didactic games. *J Balt Sci Educ*. 2020;19(6):1020–33.
15. Sneider C. Learning about seasons: a guide for teachers and curriculum developers. *Astron Educ Rev*. 2011;10(1):1–23. <https://doi.org/10.3847/AER2010035>.
16. Corrochano D, Gomez-Gonçalves A, Sevilla J, Pampin-Garcia S. Ideas of high school and university students about the meaning and origin of tides. *Eureka J Teach Pop Sci*. 2017;14(2):353–66.
17. Wang J, Guan Y, Lixin W, Guan X, Cai W, Huang J, Wenjie D, Zhang B. Changing lengths of the four seasons by global warming. *Geophys Res Lett*. 2021. <https://doi.org/10.1029/2020GL091753>.
18. Aseeva OM. Modeling as a method of understanding the surrounding reality. *Young Res*. 2021;6(348):403–4.
19. Ananishnev VM. Modeling in the field of education. *Syst Psychol Sociol*. 2010;1(2):67–84.
20. Taylor IJ, Barker M, Jones A. Promoting mental model building in astronomy education. *Int J Sci Educ*. 2010;25(10):1205–25. <https://doi.org/10.1080/0950069022000017270a>.
21. Yun E. Review of trends in physics education research using topic modeling. *J Baltic Sci Educ*. 2020;19(3):388–400. <https://doi.org/10.33225/jbse/20.19.388>.
22. López-Vargas O, Ibáñez-Ibáñez J, Racines-Prada O. Students' metacognition and cognitive style and their effect on cognitive load and learning achievement. *J Educ Technol Soc*. 2017;20(3):145–57. <https://doi.org/10.1177/1365480217704263>.
23. Cho E, Kim C, Choe S. An investigation into the secondary science teachers' perception on scientific models and modeling. *J Korean Assoc Sci Educ*. 2017;37(5):859–77.
24. Morse JM. Principles of mixed methods and multimethod research design. In: Tashakkori A, Teddlie C, editors. *Handbook of mixed methods in social and behavioral research*. Thousand Oaks: SAGE; 2003. p. 189–208.
25. Sönmez V, Alacapınar FG. *Illustrated scientific research methods*. Ankara: Anı; 2013.
26. Tekin H. *Measurement and evaluation in education* (19th edition). Ankara: Yargı Publishing; 2007.
27. Kalaycı Ş. *SPSS applied multivariate statistical techniques*. Ankara: Asil Publishing Distribution; 2008.
28. Türk C. *The effectiveness of teaching astronomy with models*. [Doctoral Thesis]. Ondokuz Mayıs University; 2015.
29. Senemoğlu N. Ensuring retention and enhancing transfer. In: *Development, learning, and teaching: from theory to practice*. Gazi Publishing House; 2004.
30. Schwarz CV, Reiser BJ, Davis EA, Kenyon L, Acher A, Fortus D, Schwartz Y, Hug B, Krajcik J. Developing a learning progression for scientific modeling: making scientific modeling accessible and meaningful for learners. *J Res Sci Teach*. 2009;46(6):632–54. <https://doi.org/10.1002/tea.20311>.
31. Chamizo JA. A new definition of models and modeling in chemistry' teaching. *Sci Educ*. 2013;22(7):1613–32. <https://doi.org/10.1007/s11191-011-9407-7>.
32. Gilbert JK. *Visualization in science education*. New York: Springer Press; 2005.
33. Ha JH, Lee HJ, Kang SJ. Perception of science high school students on modeling activity. *J Gifted Talent Educ*. 2009;19(1):184–202. <https://doi.org/10.1007/1-4020-3613-2>.
34. Şahin A, Yılmaz E. *Retention in education: a study on information recall and the learning process*. Istanbul: Eğitim Publishing House; 2016.
35. Özçelik DA. *Measurement and evaluation in schools: teacher's handbook*. Ankara: Pegem Academy; 2010.
36. Karip E. *Measurement and evaluation*. Ankara: Pegem Academy; 2009.
37. Gönen M, Katrancı M, Uygun M, Uçuş Ş. An examination of children's books for primary school students in terms of content, illustration, and physical features. *Educ Sci*. 2011;36(160):250–65.
38. Büyüköztürk Ş. *Experimental designs: pre-test-post-test control group, design, and data analysis*. Ankara: Pegem Akademi; 2011.
39. Punch KF. *Introduction to social research: quantitative and qualitative approaches*. 3rd ed. Thousand Oaks: SAGE Publications; 2014.
40. Jamieson S. Likert scales: how to (ab)use them. *Med Educ*. 2004;38(12):1217–8. <https://doi.org/10.1111/j.1365-2929.2004.02012.x>.
41. Patton MQ. *Qualitative methods and evaluation methods*. Thousand Oaks: Sage Publications; 2002.
42. Miles BM, Huberman AM, Saldana J. *Qualitative data analysis: a methods sourcebook*. Thousand Oaks: Sage Publications; 2014.
43. Şimşek H. ve Yıldırım, A. *Qualitative Research Methods in the Social Sciences*. Ankara: Seçkin; 2011.
44. Danaia L, McKinnon DH. Common alternative astronomical conceptions encountered in junior secondary science classes: why is this so? *Astron Educ Rev*. 2007;6(2):32–53. <https://doi.org/10.3847/AER2007017>.
45. Frede V. The seasons explained by refutational modeling activities. *Astron Educ Rev*. 2008;7(1):44–56. <https://doi.org/10.3847/AER2008005>.
46. Kavanagh C, Sneider C. Learning about gravity Part II. Trajectories and orbits. *Astron Educ Rev*. 2006;5(2):53–102. <https://doi.org/10.3847/AER2006019>.
47. Tsai C, Chang C. Lasting effects of instruction guided by the conflict map: experimental study of learning about the causes of seasons. *J Res Sci Teach*. 2005;42(10):1089–111. <https://doi.org/10.1002/tea.10039>.
48. Chen Y, Lin J. The impact of model-based learning on students' conceptual change in science education. *J Sci Educ Technol*. 2022;31(2):234–48. <https://doi.org/10.1007/s10956-021-09950-8>.
49. Soliman H, Moorthy S. Effectiveness of models in enhancing students' understanding of complex physical phenomena. *Int J Sci Educ*. 2021;43(4):674–89. <https://doi.org/10.1080/09500693.2020.1867391>.

50. Sánchez M, González M, García J. Addressing misconceptions in the teaching of seasonal changes: a model-based approach. *J Educ Res*. 2023;116(3):215–27. <https://doi.org/10.1080/00220671.2022.2093476>.
51. Zhang L, Liu X. Overcoming students' difficulties in understanding the Earth's tilt and orbit: Insights from model-based interventions. *Educ Res Rev*. 2021;36:100–15. <https://doi.org/10.1016/j.edurev.2021.100415>.
52. Karataş M, Yıldız M. Model-based learning strategies: enhancing students' conceptual understanding in science. *Sci Educ Int*. 2022;33(1):45–59. <https://doi.org/10.3389/feduc.2022.852580>.
53. Nair S, Kumar P, Rahman M. Models as educational tools: the role of visual and physical models in teaching complex scientific concepts. *Teach Teach Educ*. 2023. p. 112.
54. Patel R, Sharma A. Utilizing physical models to improve understanding of Earth's processes and phenomena. *J Sci Teacher Educ*. 2023;34(2):176–91. <https://doi.org/10.1080/1046560X.2022.2073435>.
55. Galano S, Çolantonio A, Leccia S, Marzoli I, Puddu E, Testa I. Developing the use of visual representations to explain basic astronomy phenomena. *Phys Rev Phys Educ Res*. 2018;14(1):1–30. <https://doi.org/10.1103/PhysRevPhysEducRes.14.010145>.
56. Trumper R. Teaching future teachers basic astronomy concepts—seasonal changes—at a time of reform in science education. *J Res Sci Teach*. 2006;43(9):879–906. <https://doi.org/10.1002/tea.20138>.
57. Children's misconceptions about weather: a review of the literature. *School Science and Mathematics*. 2000;102(5):202–215. <https://doi.org/10.1111/j.1949-8594.2002.tb18143.x>.
58. Akbulut Ö. Current approaches in reporting statistical significance in scientific research: myths and realities. *Int J East Mediterr Agric Res*. 2022;5(1):1–9.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.