Enhancing Scientific Practices In Primary Education Through A Meteorology-Oriented STEM Education Program: A Case Study

Achilleas Mandrikas, Constantina Stefanidou, Constantine Skordoulis
National and Kapodistrian University of Athens

Summary

A STEM education program entitled *Come rain or shine* implemented in a primary rural school in southern Greece as part of the *Diffusion of STEM (DI-STEM)* project and the results of its implementation are presented in this paper. The educational program deepened in weather education and intended to develop eight scientific practices for primary students proposed by the NGSS. Students' pretest and posttest questionnaires revealed difficulties in adopting meteorological vocabulary and relative scientific practices through weather measurements in their local environment. Students' answers indicate a variety in their conceptual progress depending on the scientific practice being investigated. They showed great progress in *analyzing & interpreting data, and using mathematics*, moderate progress in *developing & using models, and evaluating & communicating data)*, limited progress in *asking questions & defining problems,* and *planning & carrying out investigations* and a slight setback in *constructing explanations and engaging in argument from evidence*. Possible explanations and relative teaching implications for successful STEM education in primary education are discussed.

Keywords: STEM education, primary education, scientific practices

Introduction to scientific practices in the context of STEM education

The term *STEM education* integrates teaching and learning in the fields of science, technology, engineering, and mathematics across all grade levels in both formal and informal classroom education (Kelley & Knowles, 2016). In the last two decades STEM education has been proposed as an educational reform of science education worldwide (National Research Council, 2011; Bybee, 2013).

Since 1916, *inquiry* has been proposed by Dewey as a student-centered process in education (Barrow 2006). Since then, inquiry-based learning has been proposed as one of the most effective teaching approaches in science education (Minstrell & Van Zee, 2000). It is widely accepted that through *inquiry* students learn to solve problems of everyday life, are exercised in creative reasoning, sharpen their critical abilities, improve their ability of collaboration and communication, thus they are prepared for life (Tytler & White, 2019).

During the last decade, and specifically after the publication of the *US Framework for K-12 Science Education* (National Research Council, 2012) and the *Next Generation Science Standards* (NGSS) (National Research Council, 2013), the term *inquiry* has gradually been replaced by eight *Scientific and Engineering Practices*, which are essential for K-12 science and engineering curriculum: 1. Asking questions and defining problems, 2. Planning and carrying out investigations, 3. Developing and using models, 4. Analyzing and interpreting data, 5. Using mathematics, 6. Constructing explanations, 7. Engaging in argument from evidence, 8. Evaluating and communicating information.

These scientific practices refine *inquiry* and are proposed to be measured as students' competencies for life at the end of secondary education (National Research Council, 2012, p. 49). In this frame, a debate concerning the use of the terms *inquiry* and *science practices* has been emerged (see for example Lederman & Lederman, 2014; Tytler & White, 2019). Concerning the debate about the use of the terms *practices* and *skills* the *Framework for K-12 Science Education* makes clear that the term *practices* in scientific investigation includes not only *skills* but also *knowledge* (National Research Council, 2012, p. 30).

Moreover, *scientific practices* have been associated with many pedagogical and didactical terms in the literature of science education: problem-based approach (Kang et al., 2019), authentic experience, real-world issues (Campbell & Oh, 2015; Martín-Páez et al., 2019), hands-on activities (Kang et al., 2019; Anand & Dogan, 2021). In addition, many researchers believe that active participation through *scientific practices* reinforces students' critical thinking (Kelley & Knowles, 2016).

Nevertheless, research concerning the development of students' eight scientific practices proposed by NGSS is limited. Reiser et al. (2012) examined in detail scientific practices of explanation and argumentation and concluded that these two practices depend on each other; it means that engagement in argumentation is necessary to practice explanation construction (p. 6). In another research, *Engaging in Argumentation* was the most successfully incorporated scientific practice by secondary teachers in science class (Kang et al., 2019). Other researchers investigated how curriculum materials could support teachers to engage students in scientific practices proposed by the NGSS, but they were limited on four scientific practices (Arias et al., 2016). Mandrikas et al. (2023) investigated how a STEM educational program could enhance all eight NGSS scientific practices.

Our research investigates the development of the eight scientific practices proposed by the NGSS through the implementation of a STEM educational program in an elementary school. The implemented program entitled *Come rain or shine* deepened on weather education, an issue of everyday life, but also with environmental implications, in line with the need to promote sustainability (National Research Council, 2013; Feinstein & Kirchgasler, 2014).

Methodology

Research question

The research question of the present research is the following: to what extent the STEM educational program entitled *Come rain or shine* contributes to developing primary students' scientific practices. Specifically, the research is analyzed regarding: 1. Asking questions and defining problems 2. Planning and carrying out investigations 3. Developing and using models 4. Analyzing and interpreting data 5. Using mathematics 6. Constructing explanations 7. Engaging in argument from evidence 8. Evaluating and communicating information

Sample

The research sample consists of 60 primary education students; 40 of them are 11-year-old students (5th Grade) and 20 of them are 12-year-old students (6th Grade). They all attended the 2nd Elementary School of Messini, a rural area in southern Greece. Students had not any previous experience in STEM education. The school under consideration was included in the DI-STEM project as a *Research, Innovation and Dissemination Hub* giving the opportunity to teachers to be trained in STEM education by experts and implementing some STEM activities in their classrooms.

The STEM education program titled *Come rain or shine* was implemented in all three classrooms. The program included the construction of a weathervane, a compass and a rain gauge by everyday materials, taking measurements and finding the average of temperature, wind direction and rainfall for a month using a digital weather station as well as comparison to data from previous decades.

Data collection

Data collection took place in March and April 2023. Primary students completed the same pretest and posttest before and after the implementation of the educational program, under teachers' supervision. The questionnaire consisted of 10 questions: 7 open-ended questions, 2 multiple choice questions and 1 question based on a sketch/painting. The questionnaire was structured in accordance with the eight scientific practices proposed by the NGSS. Regarding the validity of the questionnaires, it is provided by the fact that all questions are related to eight scientific practices under consideration (content validity) and vice versa, the questionnaires included all scientific practices as they are analyzed in the NGSS. Moreover, the questionnaires were thoroughly tested by two experts in science education, experienced schoolteachers. They both agreed on the content validity of all items and in case of disagreement a third expert suggested corrections or modifications (Polit & Beck, 2006).

Data analysis

Regarding the open-ended questions, qualitative content analysis method was used to analyze the data (Mayring, 2015). The answers have been sorted in three categories named *scientifically acceptable, partially scientifically acceptable* and *unacceptable* (Shepardson et al., 2014). Criteria for this classification were the scientific explanation provided and the accuracy of the text, given the young age of the students. The range between different percentages recorded in pre-post questionnaires has been estimated and discussed. Regarding the multiple-choice questions, students' answers were categorized according to the pre-determined answers and recorded in Table 1 and Table 2. Regarding the question based on a sketch/painting, students' answers were categorized according to the content of the painting (e.g. one or more instruments)

Results & Discussion

Asking questions & defining problems

Students were asked to identify scientific and non-scientific questions by distinguishing testable and non-testable questions among five given statements (Table 1). According to pretest and posttest it seems that students showed limited progress in identifying the testable questions (statements 1, 2, 4) and the non-testable ones (statements 3, 5).

It seems that *Come rain or shine* has slightly reinforced students' ability to recognize testable questions and thus to ask questions that can be investigated. It seems that engagement in scientific investigations is a complicated procedure, which requires multiple kinds of knowledge, a fact also ascertained by other researchers (Crawford & Capps, 2018).

1.It rains more on the islands		2. Last year it rained more than the year before		3. When the birds fly close to the ground then it will rain		4. Yesterday it rained as much as it hasn't rained all winter		5. It didn't rain at all this year compared to last year because of climate change	
PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
42%	49%	53%	60%	48%	45%	33%	31%	67%	49%

Table 1. Percentages of students' correct answers about scientific and non-scientific questions

Planning & carrying out investigations

Students' ability to plan and carry out investigations has been checked through three questions. The first question asked students to produce data to serve as the basis for evidence for an explanation. Specifically, the students were asked if there are differences in rainfall between different cities in Greece (Table 2). Before the implementation of *Come rain or shine* half of the students (50%) believed that there is *moderate difference* and 23% *big difference*, while after the implementation findings revealed limited improvement (corresponding percentages reached 53% and 40%).

The second question asked students to plan and investigate using fair tests in which variables are controlled. Specifically, the students were asked to check how much rain might fall in the next week (Table 3). Before STEM activities none of the answers was scientifically acceptable and 42% were partially scientifically acceptable. In this category, we included answers that referred to an instrument (*It can be tested by rain gauges*, 6th Grade, student (st. 20) or a weather forecast (*I can see the weather on meteo.gr,* 5th Grade, st. 15), but they did not imply some kind of comparison. In addition, most of the answers (58%) were unacceptable including blank answers and tautologies (*It means that it will rain so much as never before in this winter,* 5th Grade, st. 4). After the implementation of *Come rain or shine* a few scientifically acceptable answers appeared (5%) (*To test this statement we must record the amount of rainfall during the whole winter and to compare it with that of this week*, 6th Grade, st. 10). Partially scientifically acceptable answers increased up to 55% either referring to an instrument *(It will rain a lot and we will see with the rain gauge how many millimeters*, 6th Grade, st. 14) or to a weather forecast (*We can take indications from weather stations*, 5th Grade, st. 13). Unacceptable answers decreased to 40%, and some of them revealed confusion between weather and climate (*I could do it, because the climate of my area does not have big changes, 5th Grade, st. 3*) or included some reference to climate change (*This phrase can be checked to have climate change, 5th Grade, st. 8*).

No difference		Small difference		Moderate difference		Big difference		Very big difference	
PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
8%	2%	13%	4%	50%	53%	23%	40%	5%	2%

Table 2. Percentages of students' answers about producing data for evidence for an explanation: Maria who hates the rain is trying to decide which region of Greece to move to study. Do you think there will be a difference in rain from city to city?

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unacceptable	scientifically acceptable	partially scientifically acceptable	unacceptable
0%	42%	58%	5%	55%	40%

Table 3. Percentages of students' answers about using fair tests: How do you think the sentence "Next week it will rain as much as it hasn't rained all winter" can be tested?

The third question asked students to plan and investigate using appropriate methods and tools for collecting data. Specifically, the students were asked to choose suitable objects and make a measuring instrument, device, experimental set-up, or procedure to check the rainfall and to depict it with a sketch/painting (Table 4). Even before the implementation of STEM activities many students (43%) could give scientifically acceptable answers, most of which clearly referred to the construction or the use of a rain gauge (Figure 1). Partially scientifically acceptable answers (45%) referred to meteorological instruments, to general weather measurements *(With a measuring scale of rainfall,* 5th Grade, st. 11) or to the weather forecast advice besides the rain gauge construction, while the rest 12% were considered as unacceptable answers. After the implementation of *Come rain or shine* scientifically acceptable answers increased up to 60% and some of them included the description of the scientific procedure of measurement (Figure 2), while partially acceptable decreased to 31% and unacceptable answers slightly

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unacceptable	scientifically acceptable	partially scientifically acceptable	unacceptable
43%	45%	12%	60%	31%	9%

Table 4: Percentages of students' answers about methods and tools for collecting data: Make a measuring instrument, a device, an experimental set-up or a procedure, so you can check if next week it will rain as much as it didn't rain all winter.

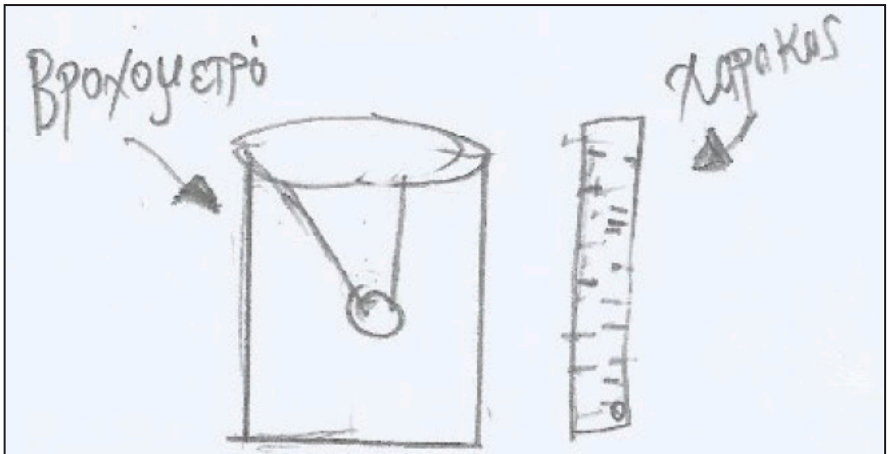

Figure 1. "A rain gauge with a ruler" (6th Grade, st. 2, pretest)

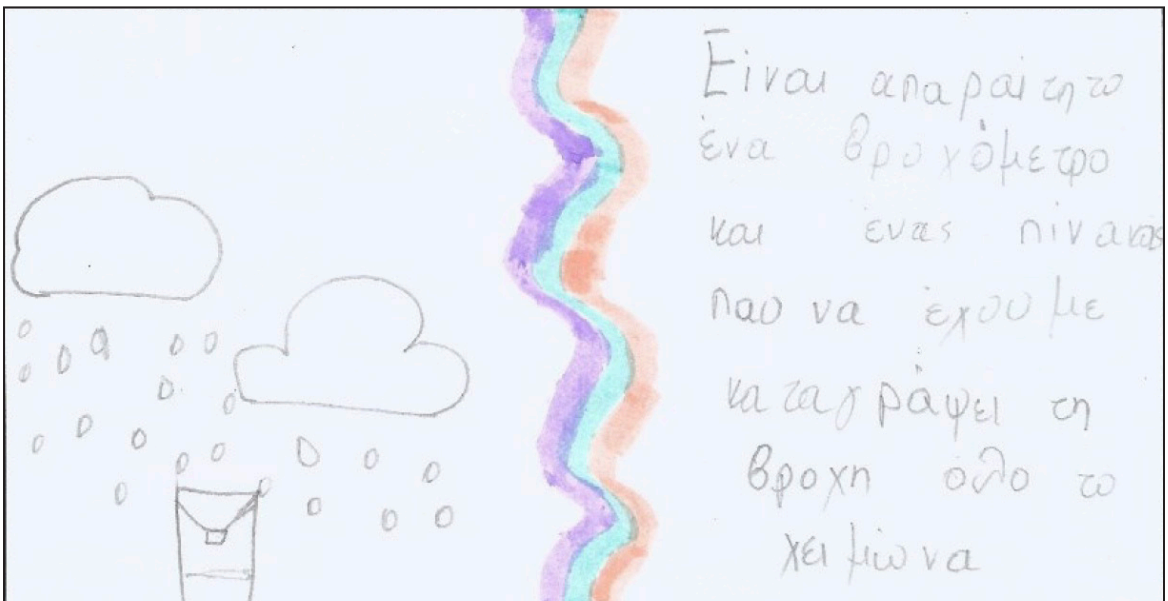

Figure 2. According to the translated text "A rain gauge is necessary and also a table on which we have recorded the rainfall for the whole winter" (5th Grade, st. 12, posttest)

decreased to 9%.

Summing up the results of the three questions we conclude that students made little progress in planning and carrying out an investigation. Although they seem able to produce data to serve as the basis for evidence for an explanation and to use appropriate methods and tools for collecting data, they are not ready to describe and use a fair test including appropriate observations, accurate measurements, finding of an average, comparisons between data recordings and correlation of variables. Our findings are compatible with other research according to which students can engage in scientific observations and scientific investigations (Metz, 2011) but at the same time face several challenges when engaging in these science practices (Eberbach & Crowley, 2009). More teacher guidance is proposed to face these challenges as well (Arias et al., 2016). Concerning research in STEM education, our findings are like Kang et al. (2019) about active students' enactment in investigation, which constitutes an authentic practice introductive to problem-solving process, as proposed by Crippen & Antonenko (2018).

Developing & using models

Concerning students' ability to develop and use models, students were asked if they could approximately predict the March temperature in Messini if weather data from previous years were provided (Table 5). Before the implementation of STEM activities, only 10% of the students gave scientifically acceptable answers (*Yes, because of the climate that is repeated with almost the same conditions every year,* 5th Grade, st. 1). A significant 43% declared that they can predict March temperature, but without any explanation or procedure (*I think I can predict March temperature*, 6th Grade, st. 1). Most of the students (47%) did not give scientifically acceptable answers, including blank answers, comparison of inappropriate variables *(I can see the temperature of this year,* 5th Grade, st. 10), some kind of connection with climate change and confusion between weather and climate (*No, I could not predict, because the weather changes,* 5th Grade, st. 11). After implementing Come rain or shine scientifically acceptable answers slightly increased up to 13% (*Yes, I could predict, because temperature values are usually the same or similar every year,* 5th Grade, st. 2), unacceptable answers decreased to 20% and most students (67%) gave partially scientifically acceptable answers.

In this question, it was expected that students would invoke the repeating weather pattern or the function of local climate as a model of explanation. Our findings revealed that students made moderate progress in using such a model, while they are probably convinced that March temperature will be approximately the same as last March, but they cannot adequately explain this pattern. However, even this moderate students' progress in developing and using models indicates that *Come rain or shine* has quite successfully transformed the scientific practice of modeling into students' experience, as proposed by other researchers (Campbell & Oh, 2015).

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unacceptable	scientifically acceptable	partially scientifically acceptable	Unacceptable
10%	43%	47%	13%	67%	20%

Table 5. Percentages of students' answers about developing & using models: Could you approximately predict the March temperature in Messini if you had weather data from previous years?

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unacceptable	scientific ally acceptabl e	partially scientifically acceptable	unacceptable
2%	27%	71%	14%	44%	42%

Table 6a. Percentages of students' answers about analyzing and interpreting data: Suppose that you have collected, using thermometer, daily temperature data of your area during March. What conclusions can you draw?

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unacceptable	scientifically acceptable	partially scientifically acceptable	unacceptable
8%	15%	77%	24%	36%	40%

Table 6b. Percentages of students' answers about analyzing and interpreting data: What other weather data would you need to be able to draw even more conclusions?

Analyzing and interpreting data

Concerning students' ability to analyze and interpret data, students were asked to answer a question in two parts. In the first part (Table 6a), students were asked for drawing conclusions after collecting daily temperatures in March. Before STEM activities 71% of the students could not give scientifically acceptable answers. As partially scientifically acceptable answers (27%) were counted these referring to general estimation of temperature (*I can tell how hot or cold it is*, 6th Grade, st. 15), these referring to continuous increase or decrease in temperature (*Every day and every hour temperature increase or decrease*, 5th Grade, st. 12) and these referring to daily temperature changes (*We can see how temperature changes every day*, 6th Grade, st. 11). After the implementation, scientifically acceptable answers became 14% of the total (*I concluded that March is a warm month with average temperature of 14,8 oC*, 5th Grade, st. 5). Partially scientifically acceptable answers increased up to 44% including those implying data comparison (*We can conclude that March of this year is warmer or colder than previous months of March*, 6th Grades, st. 8), those referring to general temperature estimation (*We can infer how warm or cold the month is*, 5th Grade, st. 3) and those referring to daily temperature changes (*Temperature in March rarely is the same*, 6th Grade, st. 10). Finally, answers that were not scientifically acceptable decreased to 42% and distributed to the same categories as before the implementation of STEM activities. According to these findings, students showed significant progress in analyzing and interpreting data (Table 6a).

In the second part (Table 6b) students were asked what other weather data they would need to draw more conclusions. Before STEM activities, 77% of the students gave scientifically unaccepted answers, such as I don't know answers and answers that denied other data (*No, we don't need other data*, 6th Grade, st. 1). As partially scientifically acceptable answers (15%) were counted answers referring to a weather station (*To find more data we put in external location an improvised weather station*, 5th Grade, st. 7), to the help of meteorologists (We would need ameteorologist to help us, 5th Grade, st. 11), to comparisons with previous years (*We can take last year's records and tell what has changed*, 6th Grade, st. 8) and these referring one or two weather parameters except temperature. After implementing *Come rain or shine* scientifically acceptable answers increased from 8% up to 24% with references to more than two weather parameters except temperature (*We should have measured humidity, rainfall, wind direction and other parameters*, 5th Grade, st. 2) or to meteorological instruments (*We will need a rain gauge, wind vane and a digital weather station*, 5th Grade, st. 8). In addition, partially scientifically acceptable answers increased from 15% to 36% with references to one or two weather parameters except temperature (*I would need the* rainfall, the humidity, 5th Grade, st. 14). Finally, unaccepted answers decreased from 77% to 40% and distributed to the same categories as before the implementation of STEM activities. The findings revealed significant progress in analyzing data for making sense of phenomena (Table 6b).

Combining findings from Table 6a and Table 6b, it is concluded that students showed noticeable progress in analyzing and interpreting data to make sense of phenomena and in using logical reasoning. Students' ability to measure quantities by using scientific instruments, to organize simple data sets, to compare data, and to reveal patterns that suggest relationships have been improved. Our findings are in line with Kang et al. (2018), who found that analyzing and interpreting data was the most successfully implemented practice by students and teachers.

Using mathematics

Concerning students' ability to use mathematics they were asked to answer a question in two parts. In the first part (Table 7a) students were asked to identify which data they consider important for estimating the weather. Before the implementation of STEM activities half of the students (50%) gave scientifically unaccepted answers. These included blank answers, answers indicating confusion between weather data and weather instruments (*The elements are thermometer and wind vane*, 5th Grade, st. 9), answers indicating confusion between weather and climate (*It is important to measure the temperature of the location and the climate of the location*, 5th Grade, st. 11) and answers referred only in weather forecast. As partially accepted were counted answers including two weather parameters (20%) (*How much cold or hot it is and how much rain or snow will fall*, 5th Grade, st. 3) and as scientifically acceptable answers including three weather parameters (30%) (*Important elements for weather estimation are temperature, wind direction and rainfal*l, 5th Grade, st. 2). After the implementation of Come rain or shine scientifically acceptable answers increased to 67%, which means that most of the students could successfully report at least three meteorological parameters. Partially acceptable answers diminished to 7% and unaccepted answers were limited to 26% either indicating confusion between parameters and corresponding instruments (*The important data are a thermometer, a wind vane, a rain gauge etc*, 5th Grade, st. 2) or indicating confusion between weather and climate (*The data from previous years*. . . , 5th Grade, st. 3).

In the second part (Table 7b) students were asked to decide whether all data can be accurately measured or some of them may be descriptive. Before STEM activities 85% of the students gave scientifically unaccepted answers, such as blank answers, and answers that revealed some misconceptions. As partially scientifically acceptable

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unacceptable	scientifically acceptable	partially scientifically acceptable	unacceptable
30%	20%	50%	67%	7%	26%

Table 7a. Percentages of students' answers about using the appropriate object or tool: Which data do you think are important for estimating the weather?

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	unacceptable	scientifically acceptable	partially scientifically acceptable	unacceptable
2%	13%	85%	31%	9%	60%

Table 7b. Percentages of students' answers about recognizing qualitative and quantitative data: In your opinion, can everything be accurately "measured" or are there some data that are more descriptive?

(13%) were counted the answers that recognized descriptive data, but without giving any examples (*No, all data cannot be accurately measured, because there are some that they are more descriptive*, 5th Grade, st. 1). After the implementation, scientifically accepted answers increased from 2% to 31% (*There are some data more descriptive like sunshine, winds, the ripple of the sea*, 5th Grade, st. 8) and unaccepted answers decreased from 85% to 60%. Finally, partially scientifically acceptable answers decreased from 13% to 9% mostly recognizing the existence of descriptive data but without giving any specific example (*There are some data that are more descriptive*, 5th Grade, st. 15).

In general, participants of the present study showed noticeable progress in using mathematics in the sense of choosing the appropriate instrument for corresponding measurements and discriminating between qualitative and quantitative data. Students' skills in using mathematics unfolds more obviously in the case of measuring, estimating, and comparing parameters to address scientific and engineering questions and refine problems. This finding is in contrast with these of Kang et al. (2018) who recorded low score in the scientific practice of using mathematics.

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unacceptabl e	scientificall y acceptable	partially scientifically acceptable	Unacceptabl e
12%	13%	75%	7%	9%	84%

Table 8: Percentages of students' answers about constructing explanations: Suppose you have placed two thermometers in two different places of the school, for example one in the yard and the other on the roof. Can you explain the difference in the two values?

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unaccepta ble	scientificall y acceptable	partially scientifically acceptable	unaccepta ble
12%	27%	61%	4%	27%	69%

Table 9. Percentages of students' answers about engaging in argument from evidence: A friend of yours claims that the hottest month in Messini is June. What would you do to ascertain whether his opinion is correct?

PRETEST			POSTTEST		
scientifically acceptable	partially scientifically acceptable	Unacceptable	scientifically acceptable	partially scientifically acceptable	unacceptable
7%	22%	71%	29%	22%	49%

Table 10. Percentages of students' answers evaluating & communicating data: You decide to present the results of your research entitled "In April 2023 it rained more than in March 2023" by making a graph. What elements do you consider essential to have in your graph?

Constructing explanations

Concerning students' ability to construct explanations, students were asked to explain the differences between two thermometers placed in different places in the school (Table 8). Before the implementation, only 12% of the students could give a scientifically acceptable answer (*The difference between the two values is due to altitude, in higher altitude it's colder while in lower altitude it's hotter*, 5th Grade, st. 2). 13% of the students answered intuitively giving partially acceptable answers (*The higher the altitude, the lower the temperature due to the mountains*, 5th Grade, st. 7) or attributed different values on measurement errors (*One thermometer may be broken and the other is working properly*, 5th Grade, st. 3). Most of the students (75%) gave unaccepted answers. Students' answers also revealed three misconceptions: 1. There is more sunshine on the roof (*The thermometer on the roof will show higher temperature, because it is closer to the sun*, 5th Grade, st. 14), 2. The thermometer on the roof is more accurate (*Because one thermometer is lower and isn't as accurate as the other,* 6th Grade, st. 13) and 3. The temperature on the roof is higher due to winds (*The thermometer on the roof shows higher temperature, because more wind "hits" the roof than the yard,* 5th Grade, st. 6). After the implementation of *Come rain or shine* scientifically acceptable answers decreased from 12% to 7% and partially scientifically acceptable answers decreased from 13% to 9%, Moreover, the unacceptable answers increased from 75% to 84%. Among them blank answers, tautologies, and the three misconceptions appeared again.

Our findings are in contrast with the literature (Reiser et al. 2012) who considered students' engagement in explanation as effective, when such engagement relates to argumentation, arguing that developing explanatory accounts includes not only construction but also comparison and critique (p. 7). This close relationship between explanation and argumentation is also highlighted by other researchers (McNeill & Krajcik, 2012). In our case, it seems that suggestions of Reiser et al. (2012) for focusing on reasons for ideas, creating a climate that is safe for students to be wrong and asking students rich questions that have multiple plausible answers were not sufficiently cultivated.

Regarding the confirmed misconceptions, they have been recorded in the relevant literature. First, the answers mentioning that the thermometer on the roof shows higher temperature because it is closer to the sun, implies the well-known misconception that air gets heat directly from the sun and not from the absorbed and then emitted heat of the ground and of the greenhouse gases (Henriques, 2002; Lutgens & Tarbuck, 2007). Second, the answers stating that the thermometer on the roof is more accurate reveal difficulties in conducting a fair test or considering the appropriate parameters, as has been found in other research concerning various kinds of measurements including temperature measurement (Keles et al., 2010). Other researchers have already found similar students' difficulties concerning the measurement of length, area and volume (Tan Sisman & Aksu, 2016) or mass (McDonough et al., 2013). Similar errors, misunderstandings and misconceptions have been recorded in other studies further investigating measurements (Carifio & Perla, 2007; Yun et al., 2016). Third, the tendency to correlate the wind speed with temperature by associating high speed with cold wind has been also reported by other researchers (Driver et al., 1994; Lee & Butler-Songer, 2003).

Engaging in argument from evidence

Concerning students' ability to engage in argument from evidence through using data to evaluate claims about cause and effect, students were asked to describe a procedure for ascertaining the opinion that the hottest month in Messini is June (Table 9). Before the implementation of STEM activities some students (12%) could give acceptable answers referring both to daily temperature measurements and appropriate comparisons between months. 27% of the answers were partially acceptable, referring either to daily temperature measurements for a month or to some comparisons between months (*We could compare the temperature of June with other months,* 5th Grade, st. 7). Most of the students (61%) gave unacceptable answers including blank answers, simple positive answers with some explanation, simple negative answers with some explanation, answers referred to a weather forecast, answers referred to some comparison with previous years and answers implying some measurements.

After the implementation scientifically acceptable answers decreased from 12% to 4% and partially scientifically acceptable answers remained unchanged. In the contrary, the unacceptable answers increased from 61% to 69%. Among them blank answers, simple positive answers with some or without explanation *(His opinion is correct,* 5th Grade, st. 13), answers referred to a weather forecast (*We will visit meteo.gr to see if it is true*, 5th Grade, st. 10), answers referred to some comparison with previous years (*I would open meteo.gr to make comparisons with other years*, 5th Grade, st. 3), answers implying some measurements (*In June we will take a thermometer to measure the temperature,* 6th Grade, st. 4), references to instruments (*We will use a thermometer*, 6th Grades, st. 9) and answers implying elements of climate (*I would look* *at all the temperatures of the last few decades and then compare*, 6th Grade, st. 1). In general, students' ability for engaging in argument from evidence has been slightly worsened despite the implementation. Possible reasons could be related to the fact that the daily measurements of the temperature for a whole month were not accompanied by further discussion and evaluation of recorded data. Maybe students were not efficiently engaged in temperature measurements, in estimation and comparisons of data, in sum and average calculation, in cause-and-effect connection into a procedure, and in argumentation to solve a problem or to check a claim.

Our findings are in contrast with these of Reiser et al. (2012) who examined in detail scientific practices of explanation and argumentation and found that students could argue for their provided explanations with elaborate and precise way which improved the causal account (p. 11). Our findings are also incompatible with these of Kang et al. (2019), who found that *Engaging in Argumentation* was the most successfully practice incorporated by secondary teachers in their science teaching practice. However, other researchers emphasize the difficulty students face when developing scientific arguments (Zembal-Saul et al., 2013) especially when they are novices in a domain (Lee & Butler-Songer, 2003). Probable explanations for these contradictable results could be based on teachers' difficulties in supporting students to familiarize with these practices, as already has been supported by other researchers (McNeill, 2009). In fact, Arias et al. (2016) after focusing on four science practices proposed by the NGSS concluded that *All teachers struggled with supporting the construction of evidence-based claims and those without supports did not push for justification for predictions (p. 1521).*

Evaluating & communicating data through tables, diagrams, and charts

Concerning students' ability to evaluate and communicate data, students were asked to present the results of a hypothetical research called *In April 2023 it rained more than in March 2023*. *What are the necessary elements included in your graph?* using a graph (Table 10). Before the implementation most students (71%) gave scientifically unacceptable answers. Partially acceptable answers were given by 22% of the students referring only one correct element (*I have to know the rainfall of these months,* 5th Grade, st. 11), while only 7% of the students gave scientifically acceptable answers depicted in a graph (Figure 3). After the implementation scientifically acceptable answers increased from 7% to 29% (*The title of the graph, the vertical axes, the rainfall and the months*, 5th Grade, st. 6). Improvement has been recorded in unacceptable answers, which decreased from 71% to 49%.

It seems that primary students showed a moderate progress in their ability to communicate scientific and technical information in written formats, including tables, diagrams, and charts. Thus, it seems that abilities

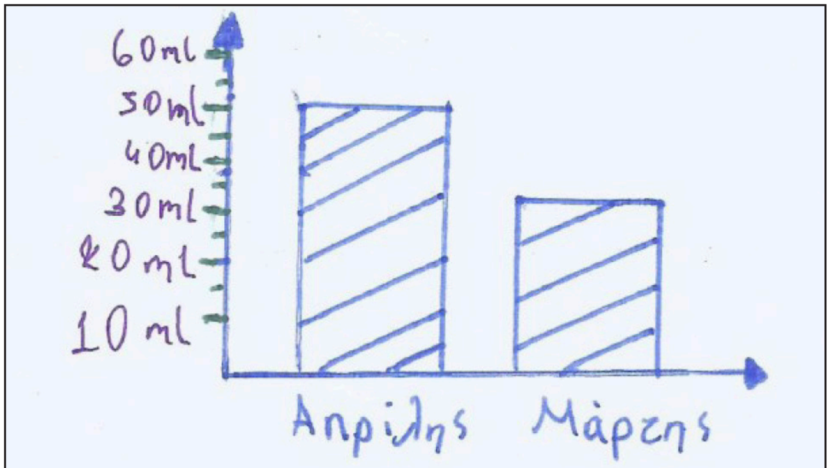

Figure 3. "In April 2023 it rained more than in March 2023" (5th Grade, student 1, pre-test)

like intuitive reading of diagrams, combining information in written text with that in corresponding diagrams, comparing data appearing in various media, using data as scientific evidence are still abilities to be further improved. Our findings seem to be more encouraging than Kang et al. (2018) who recorded very low students' scores in *Obtaining, Evaluating, and Communicating Information* as a scientific practice proposed by the NGSS and better than relative students' results in the case of orally communicating findings to classmates (Kang et al. 2019).

Conclusions and teaching implications

Our research intended to investigate the way a STEM program can contribute to the development of primary students' science practices. In this context, through *Come rain or shine* primary students were guided to construct their own meteorological instruments, to take measurements of meteorological parameters during a whole month, and to extract conclusions about their local environment based on their own recordings.

The findings of the present study vary according to the scientific practice being investigated. In particular, great progress has been recorded in a) analyzing & interpreting data, and b) using mathematics, moderate progress was recorded in a) developing & using models, and b) evaluating & communicating data, little progress was recorded in a) asking questions & defining problems, and b) planning & carrying out investigations, and a slight setback has been noted in a) constructing explanations, and b) engaging in argument from evidence. In addition, in several answers has been recorded a confusion between the terms weather and climate. Moreover, in some cases the term *climate* has been used by students as synonym of *climate change,* a term that is constantly in the news and students probably hear it without understanding it. Finally, some misconceptions have been emerged with the most well-known being that *the thermometer on the roof shows higher temperature because it is closer to the sun.*

Students' young age (5th and 6th Grade) and lack of previous experience in STEM activities could justify the unbalanced display of results depending on the scientific practice being investigated. Particularly, this is why students showed great progress in analyzing & interpreting data or using mathematics, but no progress in constructing explanations or engaging in argument from evidence. The first ones were already known in students through other subjects and were reinforced through STEM activities, but the latter ones were both new to both students and teachers, so students faced several difficulties towards them.

So, some useful suggestions for better results would be more authentic engagement of students in measurements, finding of average, discussions on recorded data, comparisons between data, and selection of appropriate time periods for comparisons. Special help for students or special instructive tools would facilitate students in constructing explanations and in engaging in argument with evidence. In general, some implications for successful STEM education in primary education would be the appropriate transformation of scientific knowledge, a concrete set of measurements, deeper estimation of data, design of authentic students' experience, and more hands-on activities.

However, there are several limitations to our research which prevent the generalization of findings and conclusions. First, students' answers were not cross checked through interviews, which could provide some clarifications in the written answers of the questionnaires. Second, the sample is limited and derived only from one elementary school. Third, the implementation was made in a rural school without any previous experience in STEM activities. Thus, suggestions for further research would be the implementation of *Come rain or shine* in a wider audience.

However, the meteorological "core" of *Come rain or shine* offered a chance to introduce STEM in primary education. Students engaged in scientific practices proposed by the NGSS and made some progress in some of them. They also had the opportunity to *experience* the stages of scientists' work, a procedure not usually occurring in typical teaching of science in class. Finally, the implementation of meteorological activities in STEM education could be an introduction to long-term measurements at the local level contributing to citizen science in the context of climate change and beyond.

Acknowledgement: The present study is supported by the HELLENIC Foundation of Research and Innovation (HFRI) in the framework of the 3rd call of action "Science and Society""Research, Innovation and Dissemination Hubs" (Project number: 01933, Title: Diffusion of STEM (DI-STEM)).

Ethics Statement: The present study is conducted as a non-interventional study with minimal or no intrusiveness, aiming at mapping students' improvement on scientific practices, without including any personal data. Ethical approval was waived under the decision of the Regional Directorate of Primary and Secondary Education (in Greece) since during the design and implementation procedures were followed to ensure ethical standards regarding the anonymity of the participants, parents' consent and password protected data storage.

References

Anand, N., & Dogan, B. (2021). Impact of informal learning environments on STEM education - Views of elementary students and their parents. *School Science and Mathematics, 121*(6), 369–377. https://doi.org/10.1111/ssm.12490

Arias, A. M., Davis, E. A., Marino, J.-C., Kademian, S. M., & Palincsar, A. S. (2016). Teachers' use of educative curriculum materials to engage students in science practices. *International Journal of Science Education, 38*(9), 1504-1526. https://doi.org/10.1080/09500693.2016.1198059

Barrow, L. H. (2006). A brief history of inquiry: From Dewey to standards. *Journal of Science Teacher Education, 17*(3), 265–278. https://doi.org/10.1007/s10972-006-9008-5

Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities*. National Science Teachers Association Press. https://static.nsta.org/pdfs/samples/PB337Xweb.pdf

Campbell, T., & Oh, P. S. (2015). Engaging Students in Modeling as an Epistemic Practice of Science: An Introduction to the Special Issue of the Journal of Science Education and Technology. *Journal of Science Education and Technology, 24*, 125–131. https://doi.org/10.1007/s10956-014-9544-2

Carifio, J., & Perla, R. J. (2007). Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. *Journal of Social Sciences, 3*(3), 106-116. http://dx.doi.org/10.3844/jssp.2007.106.116

Crawford, B. A., & Capps, D. K. (2018). Teacher Cognition of Engaging Children in Scientific Practices. In Y.J. Dori, Z. R. Mevarech, D. R. Baker (eds.), Cognition, Metacognition, and Culture in STEM Education, Innovations in Science Education and Technology 24, p. 9-32. https://doi.org/10.1007/978-3-319-66659-4_5

Crippen, K. J., & Antonenko, P. D. (2018). Designing for Collaborative Problem Solving in STEM Cyberlearning. In Y.J. Dori, Z. R. Mevarech, D. R. Baker, *Cognition, Metacognition, and Culture in STEM Education, Innovations in Science Education and Technology24*, p. 89-116. https://doi.org/10.1007/978-3-319-66659-4_5

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science: Research into children's ideas*. London: Taylor & Francis Ltd.

Eberbach, C., & Crowley, K. (2009). From everyday to scientific observation: How children learn to observe the biologist's world. *Review of Educational Research, 79*(1), 39–68. https://doi.org/10.3102/0034654308325899

Feinstein, N. W., & Kirchgasler, K. L. (2015). Sustainability in Science Education? How the Next Generation Science Standards Approach Sustainability, and Why It Matters. *Science Education, 99*(1), 121-144. https://doi.org/10.1002/sce.21137

Henriques, L. (2002). Children's ideas about weather: A review of the literature. *School Science and Mathematics, 102*(5), 202–215. https://doi.org/10.1111/j.1949-8594.2002.tb18143.x

Kang, E. J. S., Donovan, C., & McCarthy, M. J. (2018). Exploring elementary teachers' pedagogical content knowledge and confidence in implementing the NGSS science and engineering practices. *Journal of Science Teacher Education, 29*(1), 9–29. https://doi.org/10.1080/1046560X.2017.1415616

Kang, E. J. S., McCarthy, M. J., & Donovan, C. (2019). Elementary Teachers' Enactment of the NGSS Science and Engineering Practices. *Journal of Science Teacher Education, 30*(7), 788-814. https://doi.org/10.1080/1046560X.2019.1630794

Keles, O., Ertas, H., Uzun, N., & Cansiz, M. (2010). The understanding levels of preservice teachers' of basic science concepts' measurement units and devices, their misconceptions and its causes. *Procedia - Social and Behavioral Sciences, 9*, 390–394. https://doi.org/10.1016/j.sbspro.2010.12.170

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*, 11. https://doi.org/10.1186/s40594-016-0046-z

Lederman, N. G., & Lederman, J. S. (2014). The Next Generation Science Standards: Implications for Preservice and Inservice Science Teacher Education. *Journal of Science Teacher Education, 25*, 141–143. https://doi.org/10.1007/s10972-014-9382-3

Lee, H.-S. & Butler-Songer, N. (2003). Making authentic science accessible to students. *International Journal of Science Education, 25*(8), 923-948. https://doi.org/10.1080/09500690305023

Lutgens, F. K., & Tarbuck, E. J. (2007). *The Atmosphere, An Introduction to Meteorology.* 10th Ed., New Jersey, Pearson Prentice Hall, USA.

Mandrikas, A., Stefanidou, C., Kyriakou, K., &. Skordoulis, C., (2023). Scientific Practices in the Context of STEM Education: A Case Study in Primary Education. Journal of STEM Education, 24(3), 35-44. https://jstem.org/jstem/index.php/JSTEM/article/view/2663/2340 Martín-Páez, T., Aguilera, D., Perales-Palacios, F.J., & Vílchez-González, J.M. (2019). What are we talking about when we talk about STEM education? A re-view of literature. *Science Education, 103*, 799–822. https://doi.org/10.1002/sce.21522

Mayring, P. (2015). Qualitative Content Analysis: Theoretical Background and Procedures. In A. Bikner-Ahsbahs, C. Knipping & N. Presmeg (Eds.), Approaches to Qualitative Research in Mathematics Education (pp. 365–380). Dordrecht: Springer. https://link.springer.com/chapter/10.1007/978-94-017-9181-6_13

McDonough, A., Cheeseman, J., & Ferguson, S. (2013). Young childrens' emerging understandings of the measurement of mass. *Australasian Journal of Early Childhood, 38*(4), 13-20. https://doi.org/10.1177/183693911303800403

McNeill, K. L. (2009). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education, 93*(2), 233–268. https://doi.org/10.1002/sce.20294

McNeill, K.L., & Krajcik, J. (2012). *Supporting grade 5–8 students in constructing explanations in science: The claim, evidence, and reasoning framework for talk and writing*. New York: Allyn and Bacon.

Metz, K. E. (2011). Disentangling robust development constraints from the instructionally mutable: Young children's epistemic reasoning about a study of their own design. *Journal of the Learning Sciences, 20*(1), 50–110. https://doi.org/10.1080/10508406.2011.529325

Minstrell, J., & Van Zee, E. H. (eds) (2000). *Inquiring into Inquiry Learning and teaching in Science*. American Association for the Advancement of Science, Washington, DC.

National Research Council (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Washington: The National Academies Press. https://nap.nationalacademies.org/catalog/13158/successful-k-12-stem-education-identifying-effective-approaches-in-science

National Research Council (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, DC: The National Academies Press. https://doi.org/10.17226/13165

National Research Council (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18290

Reiser, B. J., Berland, L. K., & Kenyon, L. (2012). Engaging Students in the Scientific Practices of Explanation and Argumentation. *Science Scope, 35*(8), 6-11. https://static.nsta.org/ngss/resources/201204_Framework-ReiserBerlandKenyon.pdf

Polit, D.F. & Beck, C.T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health, 29*, 489-497. https://doi.org/10.1002/nur.20147

Shepardson, D. P., Roychoudhury, A., Hirsch, A., Niyogi, D., & Top, S. M. (2014). When the atmosphere warms it rains and ice melts: seventh grade students' conceptions of a climate system. *Environmental Education Research, 20*(3), 333-353. http://dx.doi.org/10.1080/13504622.2013.803037

Tan Sisman, G., & Aksu, M. (2016). A Study on Sixth Grade Students' Misconceptions and Errors in Spatial Measurement: Length, Area, and Volume. *International Journal of Science and Mathematics Education, 14*, 1293-1319. https://doi.org/10.1007/s10763-015-9642-5

Tytler, R., & White, P. (2019). Framing and Assessing Scientific Inquiry Practices. In Y.-S. Hsu and Y.-F. Yeh (eds.), *Asia-Pacific STEM Teaching Practices,* p. 173-189. https://doi.org/10.1007/978-981-15-0768-7_11

Yun, H.J., Ko, ES., & Yoo, Y.J. (2016). *Students' Misconceptions and Mistakes Related to Measurement in Statistical Investigation and Graphical Representation of Data.* In: Ben-Zvi, D., Makar, K. (eds) The Teaching and Learning of Statistics. Springer, Cham. https://doi.org/10.1007/978-3-319-23470-0_14

Achilleas Mandrikas has been a primary school teacher since 1987. He holds an MA in Science Education (2006) and a PhD in Environmental Science Education (2010), both from the National and Kapodistrian University of Athens. Among 2009-2018 he was a School Advisor in primary education for the Hellenic Ministry of Education. Among 2018-2023 he was an Educational Coordinator on Education for Sustainability in the 6th Regional Centre of Educational Planning of Attica, Greece. In parallel, he also lectures undergraduate and postgraduate students in the Department of Primary Education of the National and Kapodistrain University of Athens on environmental science and environmental education.

Constantine Skordoulis is Professor of Epistemology and Didactical Methodology of Physics in the Department of Primary Education, National and Kapodistrian University of Athens. He is the Director of the Laboratory of Science Education and Educational Technology and Head of the Postgraduate Studies Program "STEM Education and Educational Robotics Systems". He has studied Physics at the University of Kent at Canterbury, UK and worked as a visiting researcher at the Universities of Oxford and Groningen. He is Co-Editor of "Almagest / Journal for the History of Scientific Ideas" (Brepols, Brussels) and member of the Editorial Boards of international journals (Journal of Critical Education Policy Studies, Advances in Historical Studies, Journal of Interdisciplinary Social Science etc). He is an Effective Member of the International Academy of History of Science and has served as the Secretary of the Interdivisional Teaching Commission of the International Union of History and Philosophy of Science (2007-17). His research interests include History of Science, Science Education and STEM Education from a critical perspectiv

Constantina Stefanidou is a physicist who has a PhD in Science Education from University of Athens. After a long period of teaching science in secondary education, she now has a position in the Department of Education at Athens University in the field of pre-service teachers' science education. Her research and publications are on STEM Education and didactics of science, focusing on historical and philosophical perspectives as well as conceptual difficulties and their relation to model-based teaching and learning and informal science education. She participates in international conferences and science communication actions. She is a member of the European Science Education Research Association (ESERA) and the International History Philosophy and Science Teaching (IHPST) Group.

Zembal-Saul, C. L., McNeill, K. L., & Hershberger, K. (2013). *What's your evidence? Engaging K-5 children in constructing explanations in science*. Boston, MA: Pearson Education.