JOURNAL OF BALTIC SCIENCE EDUCATION
ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

# PRIMARY SCHOOL STUDENTS' DIFFICULTIES IN WRITING ARGUMENTS: IDENTIFYING CHALLENGES AND OPPORTUNITIES FOR SCIENCE TEACHING

**Tomokazu Yamamoto,**
**Shinichi Kamiyama,**
**Tatsuya Tanaka,**
**Etsuji Yamaguchi**

**Abstract.** *In science education, the improvement in students' ability to construct arguments at the primary school level has been reported. Although these studies have identified difficulties in arguments written by primary school students, they do not indicate areas that require improvement in teaching methods. This study aims to explore the possibility of improvement in early primary school students' ability to construct arguments and identify the types of difficulties encountered. Sixty-seven Japanese third-grade students (9–10 years old) were taught to write arguments as specified by Zembal-Saul et al. (2012). The students were given two writing tasks before and after the lesson. To examine the students' written arguments, each component of claim, evidence, and reasoning was scored based on a rubric. On comparing the scores of the pre-test and post-test writing tasks, it was found that 27 out of 67 students still had difficulty writing arguments during the post-test. An analysis of the students' writing revealed four types of difficulties: 'Incompleteness of components', 'Inappropriateness of components', 'Confusion between evidence and reasoning', and 'Confusion between claim and evidence'. This study offers insights pertaining to teaching implications and research recommendations.*
**Keywords:** *difficulties in writing arguments, elementary/primary school, explanation construction, small-sample quantitative study*

**Tomokazu Yamamoto**
*Hyogo University of Teacher Education 1, Japan*
**Shinichi Kamiyama**
*Osaka University of Health and Sport Sciences 2, Japan*
**Tatsuya Tanaka**
*Kobe Municipal Nadanohama Primary School 3, Japan*
**Etsuji Yamaguchi**
*Kobe University 4, Japan*

## Introduction

In knowledge construction in science classes, learners must persuade others based on evidence. To obtain agreement from others, learners need to explain their ideas logically and rationally, and argumentation is the focus of attention as a structure of explanation for this purpose. Argumentation has a structure consisting of the claim, data, warrant, backing, qualifier, and rebuttal (Toulmin, 1958), and these have been used to introduce the teaching of argumentation in science lessons (e.g., Erduran & Jiménez-Aleixandre, 2007; Osborne et al., 2004; McNeill & Krajcik, 2011; Saracaloglu et al., 2011). Based on Toulmin's Argument Pattern (TAP), analysing learners' discourse based on its components has been widely used not only in science but also in subjects such as history, social issues, and socio-scientific issues (SSI). These analyses have reported that, with appropriate support and environment, learners can construct arguments. For example, Monteira, et al. (2016) coded kindergartners' utterances with claims, raw data, evidence, and justification on whether snails can distinguish sounds and demonstrated that even kindergartners can claim observations as evidence. In addition, Pontecorvo, who has conducted several studies focusing on children's discourse, implemented a curriculum for fourth-grade primary school students in Rome to critically read historical documents and found that even young children can construct arguments based on claims, justification, concession, opposition, and counter opposition. Pontecorvo and Girardet (1993) demonstrated that autonomous interactional activities produce high levels of reasoning even in young children. Similarly, Kuhn et al.'s research, which emphasised interactional activities for primary and middle school students, revealed the outcomes of interactive interventions on social issues and SSI, allowing for discussions that consider opposing positions (Kuhn & Crowell, 2011; Kuhn et al., 2016; Iordanou et al., 2019). These previous studies indicate that explanations using arguments are essential in the construction of knowledge in science classes and that teaching in this area is emphasised from the primary school level.


https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

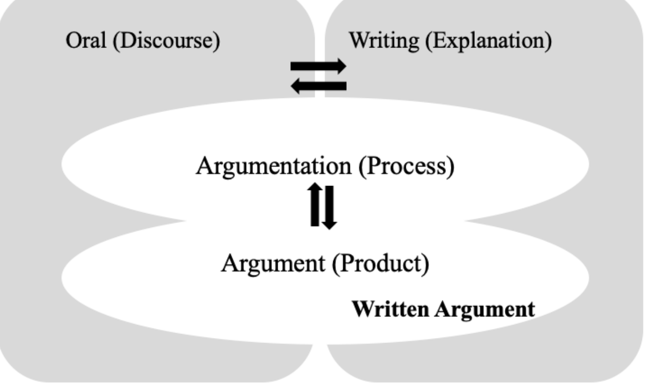
*Importance of Written Arguments*

This study deals with written arguments, not oral ones. Figure 1 illustrates the position of written arguments in the dichotomy between oral discourse and writing explanations. A written argument is an explanation that states the answer to a question. Osborne et al. (2016) defined an argument as a product and argumentation as a process and highlighted the relevance of research that focuses on students' ability to use argumentative reasoning. In terms of the reasons for focusing on written arguments, first, they demonstrated that learners could engage in simulated dialogue through writing tasks in a dialogue intra-mentally. Second, they pointed out that 'argumentation as practiced by scientists often occurs exclusively in writing, via published articles and written responses. Written arguments are often the foundation of the practice of argumentation' (Osborne et al., 2016, p. 31). In other words, the ability to construct a written argument is the most fundamental ability for facilitating argumentation and oral discourse.

Moreover, McNeill and Krajcik conducted studies on teaching strategies and the assessment of written arguments. They mentioned the benefits of having students write scientific explanations, as the activity helps students: '(1) understand science concepts, (2) develop twenty-first-century skills, (3) use evidence to support claims, (4) reason logically, (5) consider and critique alternative explanations, and (6) understand the nature of science' (McNeill & Krajcik, 2011, p.7). They see writing as an attempt to persuade others of the validity of scientific knowledge and argue that making students write scientific arguments supports a deeper understanding of content and engages students in the social scientific practice of debating and constructing knowledge claims using evidence and reasoning (McNeill, 2009). Furthermore, Zembal-Saul et al. (2012) argued that:

> *In our work in elementary classrooms, talking and writing science explanations are complementary activities. Sometimes we engage students in talking about their ideas first (e.g., predictions) in preparation for an investigation in which they will document observations in writing in their science notebooks, which will later serve as evidence for scientific claims. Other times we have students attempt to identify patterns in evidence and/or draft an initial claim in writing before gathering for a science talk in which we collectively construct claims from evidence (p. 12).*

**Figure 1**
*The Position of the Written Argument*




As for writing argumentations (the process) in science lessons, the practice and results have been reported mainly for lower-secondary school and upper-secondary school students. Clark and Sampson (2007) conducted an experiment where students had to construct principles to explain the data they had collected through an online dialogue system and then assigned the students into discussion groups based on the different principles so that they could consider and critique the various perspectives. Chin and Osborne (2010) asked eighth-grade students in London and Singapore to construct an argument based on a graph of the three-state changes of water. They then analysed the transcripts of the students' questions and discussions and conducted a practical study on the 'Questioning-Argumentation model', which focussed on questioning. Kuhn (2010) and Iordanou (2010) developed a three-step process for sixth and seventh-grade students in the US and Cyprus to present effective arguments in both scientific and non-scientific domains. They assisted the students in their argumentation through group build-


https://doi.org/10.33225/jbse/22.21.445

ing and the use of sheets. In addition, they had 16-year-olds and eleventh-grade students construct an argument with an opponent via an electronic medium and reported an increase in the use of evidence by the students to weaken the opponent (Iordanou & Constatinou, 2015).

In a study focusing on the written argument (product), Ping et al. (2020) pointed out that 'Students' lack of argumentation skills in science is a result of their misunderstanding of what constitutes argumentation in science' (p. 278). They implemented the Modified Argument-Driven Inquiry (MADI) approach—focusing on scientific process skills and using arguments in a practical context—with middle school students studying biology. After analysing the students' writings, they reported that students in the MADI group improved their skills and were able to provide more evidence in their arguments. Sandoval and Reiser (2004) conducted a practical study using the Explanation Constructor argument-supporting software as an epistemic tool for obtaining evidence-based explanations from ninth-grade American students on how low rainfall in the Galapagos Islands has affected the finch population. Moreover, Kaya (2013) found that students understood the underlying concepts of chemical equilibrium better when they constructed an argument than when they did not.

*Written Arguments by Primary School Students*

Zembal-Saul et al.'s (2012) argument framework was adopted in this study because it demonstrates how to write an argument in many primary school science classes. In their framework, an argument is a statement of explanation for a question, and the framework is mainly aimed at primary school children. Their study recommends an argument structure consisting of three elements: the claim (the answers to questions or problems), evidence (the scientific data supporting the claim), and reasoning (the articulation of why the evidence supports the claim), for beginners. For example, 'Do bush bean plants grow better in direct sunlight?' is explained by students as 'Bush bean plants grow better in direct sunlight (claim) because the plant in direct sunlight grew 16 cm, the plant with less sunlight grew 11 cm (evidence), and the plant in direct light was able to grow better (reasoning)'.

Using the three elements, the lower primary grade students were asked to write an argument using the three elements of claim, evidence, and reason: 'In a magnetic fishing game, there are limited materials that can be used as parts that can be attached to the mouth of the fish (claim), because steel wool and iron wire have been attached to magnets (evidence), and it is thought that the objects attached to magnets are iron (reasoning)'.

*Primary School Students' Difficulties in Writing Arguments*

Research involving students describing the arguments mentioned above has been conducted mainly with upper primary school students (upper fifth-grade). Zembal-Saul et al. (2012) proposed and advocated for a specific method of teaching argumentation to K-5 students. Additionally, McNeill (2011) reported that fifth-grade students' written scientific arguments improved over the school year, and the 'students' ideas aligned more closely with the scientific perspectives of explanation, argument, and evidence for science and science classrooms at the end of the school year' (p. 818). Inspired by McNeill's work, Hong et al. (2013) conducted a class in which small groups of fifth-grade students discussed controversial issues in society, such as the use of plastic products, internet ethics, and doctors' notifications to patients. According to Osborne et al.'s (2004) evaluation framework, the quality of the discussion in the experimental group was higher than in the comparison group. They noted that students enjoyed learning the TAP model, and it increased their interest in learning science, demonstrating that primary school students can benefit from learning argumentation through scientific and societal intervention. Moreover, in Japan, it has been reported that upper primary school students improved their ability to write arguments, including each component, through science lessons in which they construct arguments (e.g., Kamiyama et al., 2015; Yamamoto et al., 2013).

While the effectiveness of the teaching on how to write arguments has been reported, previous studies have demonstrated the challenges regarding the written arguments of primary school students. For example, McNeill and Krajcik (2011) reported a case in which a fifth-grade student was asked to determine the size of a hawk population when all seeds were removed from the ecosystem, but the evidence presented by the student in support of their answer was insufficient, and the student was unable to write a rationale for why the evidence supported the claim. McNeill and Krajcik (2011) also highlighted cases where the claim did not reflect the correct answer to a question, or the claim was incorrect. Similarly, when Zembal-Saul et al. (2012) asked fourth-grade students about the effect of washers on the speed of a car, the students could not adequately provide evidence or reasoning.


https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

*Issues in Previous Research*

The previous studies mentioned above remain problematic for two reasons. First, the effectiveness and difficulties of teaching written arguments have been found in some studies on upper primary school students and above, but few studies have been conducted with younger students. There is a need to ascertain whether the teaching of written arguments is effective for students learning science for the first time. Second, there is a lack of clarity on younger students' difficulties with writing arguments. It is unclear what difficulties exist for students in lower primary grades who are learning argumentation in science for the first time, even though it is likely that some students have serious difficulties. To identify the challenges and opportunities for teaching how to write arguments, it is necessary to improve the written arguments of early primary school students and identify what difficulties they have after teaching.

*Research Aim and Research Questions*

In this study, lessons were conducted using Zembal-Saul et al.'s (2012) writing scaffold to confirm the improvement of argumentation skills and focus on the students' difficulties that remained. The research questions were as follows:

(1) Can science lessons based on Zembal-Saul et al.'s writing scaffold improve the written arguments of third-grade primary school students (aged 9–10)?
(2) What types of students' difficulties emerge after learning argumentation?

## Research Methodology

*General Background*

This research was a small-sample quantitative study with students (aged 9–10) from the third grade of primary school. In this study, primary argument teaching was introduced, including claims, evidence, and reasoning, using the five writing scaffolds developed by Zembal-Saul et al. (2012). The students were given a writing task before and after the lesson. Additionally, through an analysis of the writing tasks, the students' improvements and difficulties were clarified. The science lessons were conducted in February 2019 for 12 credit hours (1 credit hour is 45 minutes), with writing tasks conducted before lessons in February 2019, and after lessons in March 2019.

*Writing Scaffolds in Argument Construction*

Table 1 illustrates the unit 'Properties of Magnets (total of 12 hours)' activities. In the first and second periods (periods means credit hours), the students played with magnetic toys to familiarise themselves with magnets. They identified the following three problems for determining the properties of magnets:

(1) In a magnetic fishing game, are there limited materials that can be used as parts that can be attached to the mouth of the fish?
(2) When making a maze in which a magnet is applied from under the paper to advance the frame on the paper, can the paper be thick?
(3) Can we make a magnetic car with an S-pole move forward?

From the third to the ninth periods, the students were asked to conduct experiments to determine the properties of magnets for a fourth learning problem, along with the three problems set up in the previous period. This problem question was 'Can a magnet made of steel make a magnetic car move forward?' to determine the phenomenon of magnetisation.

From the tenth to the twelfth period, the students reviewed the properties of magnets that they had learned in the previous period, and each student improvised a magnetic toy that used the properties of magnets. Then, they introduced their toys to the other students.


https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

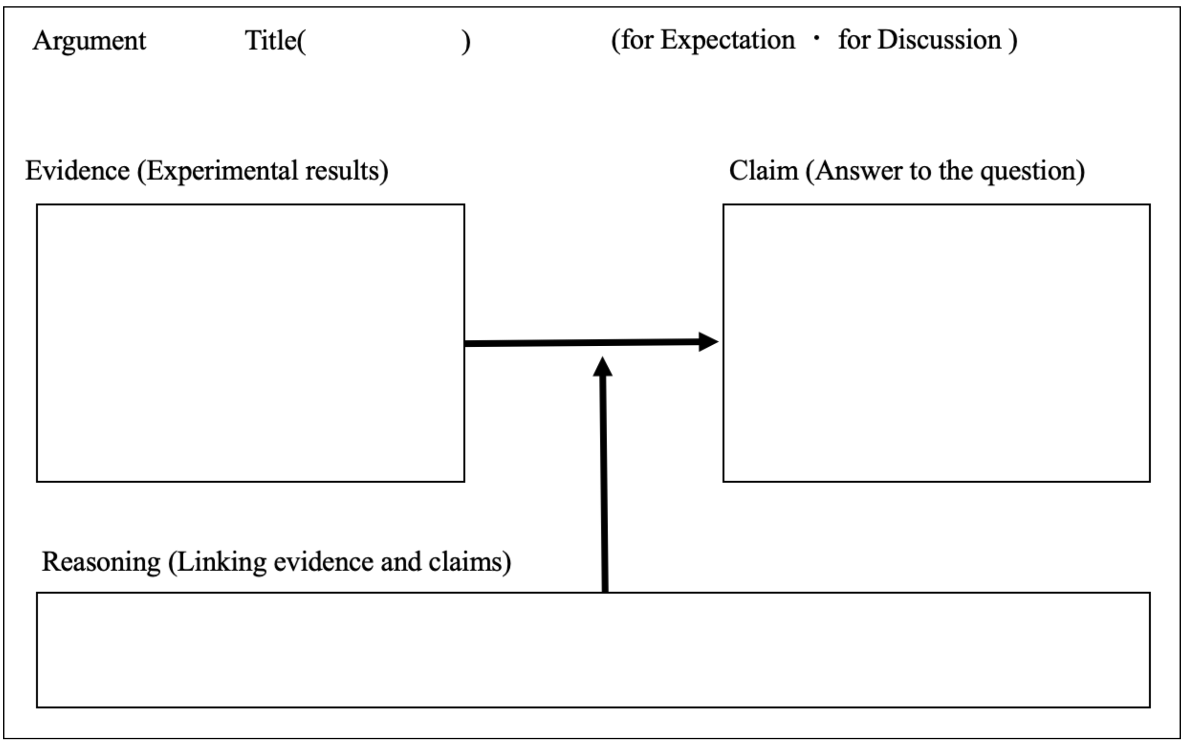

**Table 1**
*Unit Activities*

| Period | Activities |
|---|---|
| 1–2 | Play with magnetic toys to familiarise oneself with magnets. |
| 3–4 | Find out that some objects are attracted to magnets, and some are not. (Written argument) |
| 5–6 | Find out that the force of attraction works even when there is a gap between the magnet and the object. (Written argument) |
| 7–8 | Find out that different poles of a magnet attract each other, and the same poles repel each other. (Written argument) |
| 9 | Find out that rubbing a magnet against iron magnetises it. (Written argument) |
| 10–12 | Create a toy using magnets and present it to the other students. |

To help students develop the ability to construct arguments, support was provided based on the writing scaffold outlined by Zembal-Saul et al. (2012). These scaffolds are sentence starters, questions, or other prompts that provide students with hints about what to include in their argument. While designing the scaffolds, the following five features were considered:

1. **General and content support**
   'General support provides hints about the framework that could be used in any content area. Content-specific support provides hints about the specific content or task that the students are trying to answer' (Zembal-Saul et al., 2012, p. 76). As an embodiment of this idea, previous studies have used worksheets that visually illustrate the structure of an argument to assist in fitting contextual content into a general argument framework (e.g., Chin & Osborne, 2010; Çoban, 2013). The worksheet in Figure 2 was developed with reference to the 'Argument diagram' by Chin and Osborne (2010). This is a prompt about what to write in the claim, evidence, and reasoning columns. In class, content support was provided by using claim, evidence, and reason cards and writing on the board which part of the content each element corresponded to (Figure 3).

**Figure 2**
*Written Argument Worksheet*

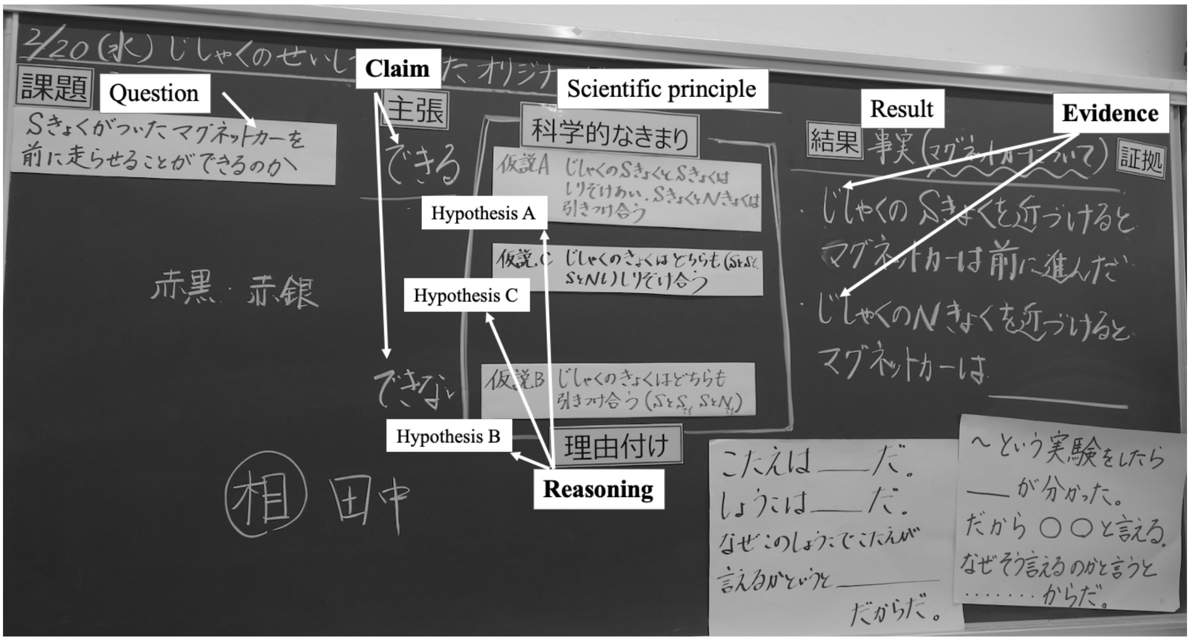




https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

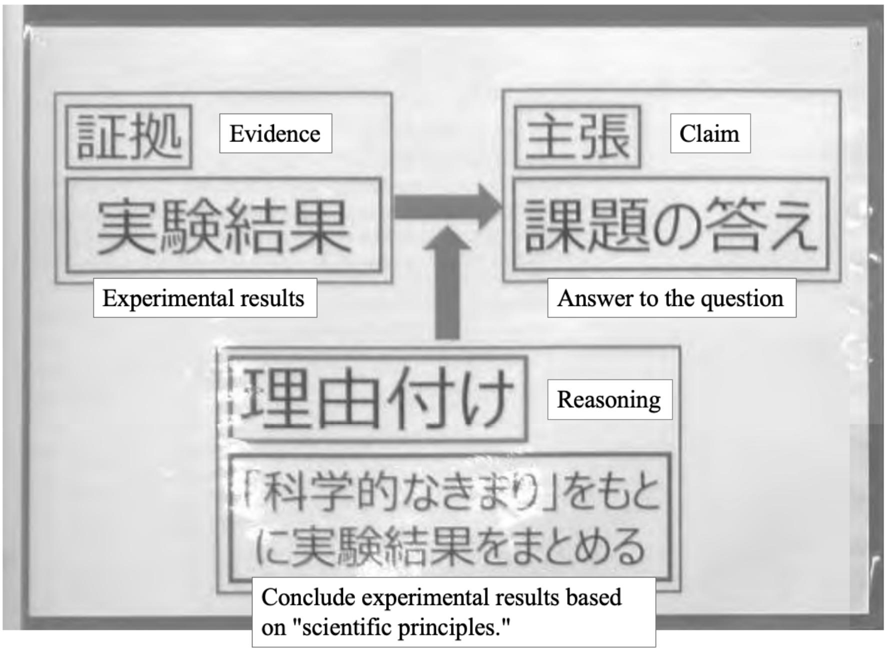
**Figure 3**
*The Claim, Evidence, and Reasoning Cards on the Board*




2. **Detail and length**
Since this was the first time the students worked on the argument structure, detailed scaffolding was conducted. The teacher gave detailed oral advice to the students as they wrote their arguments and asked them to confirm the following points: to make a 'claim' that directly addressed the question, use the experimental results in tables and bullet points as 'evidence', and use the hypothesis as 'reasoning'. In the worksheets used in the first two of the four argument writing sessions, the 'prompt' for each component was expressed in a short and simple manner, and the relationship between the components was displayed in a box so that the students understood it better through the visual representation.
3. **Fading**
'It can be important to fade the support you provide students over time to help them internalise the framework and be able to apply it to new situations' (Zembal-Saul et al., 2012, p. 79). In this study, the students were asked to write their arguments freely without using the worksheet after three of the four argument writing sessions. When the students performed the peer-evaluation of the free writing of the argument task, they used three coloured pens to colour-code and draw a line through the claim, evidence, and reasoning parts.
4. **Structure**
The worksheets supported the awareness of the structure of the argument with explanations, sentence starters, and questions. When the worksheet was presented, the prompt corresponding to each component was displayed, and sentence starters such as 'The answer is _____. The evidence is ____. The reason this answer can be given is because it is _______________.' were posted for the students to refer to during the free writing.
5. **Visual representation**
Figure 4 illustrates the model of the argument structure shown to the students that was posted on the blackboard throughout the lesson. The students were encouraged to write their arguments with an awareness of the argument structure and present evidence for the experimental results, a claim for their answers, and apply reasoning based on scientific principles.


https://doi.org/10.33225/jbse/22.21.445

**Figure 4**
*Model of the Argument Structure Shown to the Students*




*Sample Selection*

The participants were 67 third-year students (34 boys and 33 girls, aged 9–10) from a national university-affiliated school in Kobe, Japan. The students belonged to the grade in which they first learned science in primary school and had received previous teaching on argumentation in science classes. The teacher of both classes is one of the authors, who has been teaching primary school for 12 years and specialises in science education, with experience in teaching argumentation. Before the task, all students and their parents were informed that the task was for research and was unrelated to their grades. They also explained that the collected sheets would only be used for scientific, educational research, individual names will be removed during the data processing, and individuals could not be identified from the survey items. All students and their parents agreed. To ascertain the improvement in the students' ability to construct arguments through the lessons, two types of writing tasks were conducted twice, first in February 2019, before the start of the unit, and second, in March 2019, after the end of the unit. The response time allowed for each task was 15 minutes.

*Instrument and Procedures*

The two writing tasks administered before and after the class required the students to construct arguments and were related to different areas in the Course of Study as Japanese primary science standards: (1) measuring the temperature change of the ground from the domain 'Life and Earth' (hereinafter referred to as 'the ground task') and (2) measuring the weight of soil and wood in the domain 'Material and Energy' (hereinafter referred to as 'the weight task'). The students had already learned these contents, and both types of writing tasks involved presenting quantitative data from observations and experiments, as well as two options for answering the question. The students were then asked to write their answers to the questions supported by reasons, and the details of the ground task and the weight task are described below.

1. **The ground task**
   Figure 5 illustrates the ground task, in which the temperatures of two points on the ground ('Ground in A' and 'Ground in B'), measured at 9:00 am and noon, were presented on graphs. The students were then asked to describe which point was in the sun and which was in the shade and why they thought so.


https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

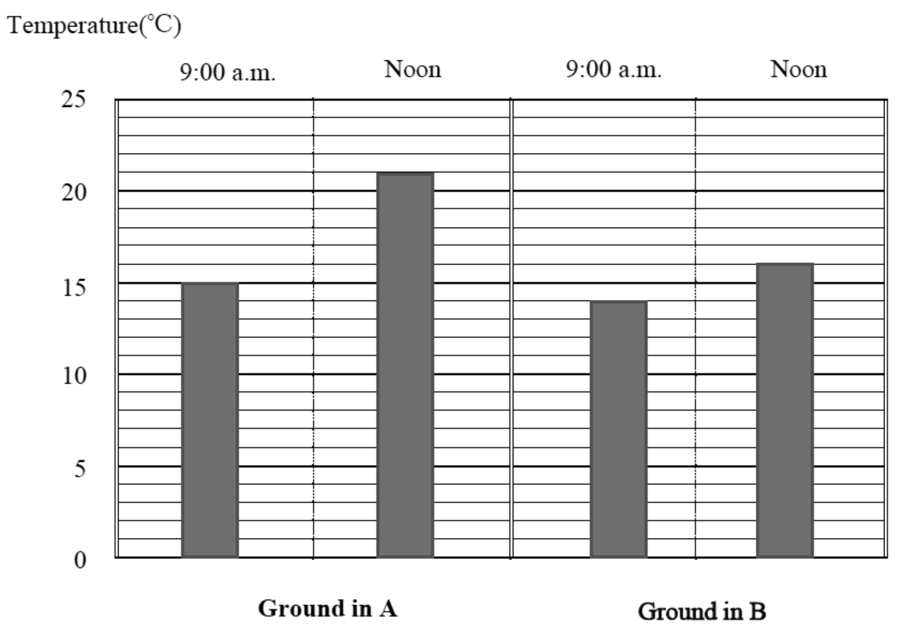
**Figure 5**
*The Ground Task*

**The Ground Task**

We measured the temperature of the ground in the sun and in the shade at the playground.

**Question**

We measured the temperature of the ground in the sun and in the shade at the playground. In the graph below, which ground (from A and B) was in the shade and in the sun? Why do you think so? Explain this scientifically.

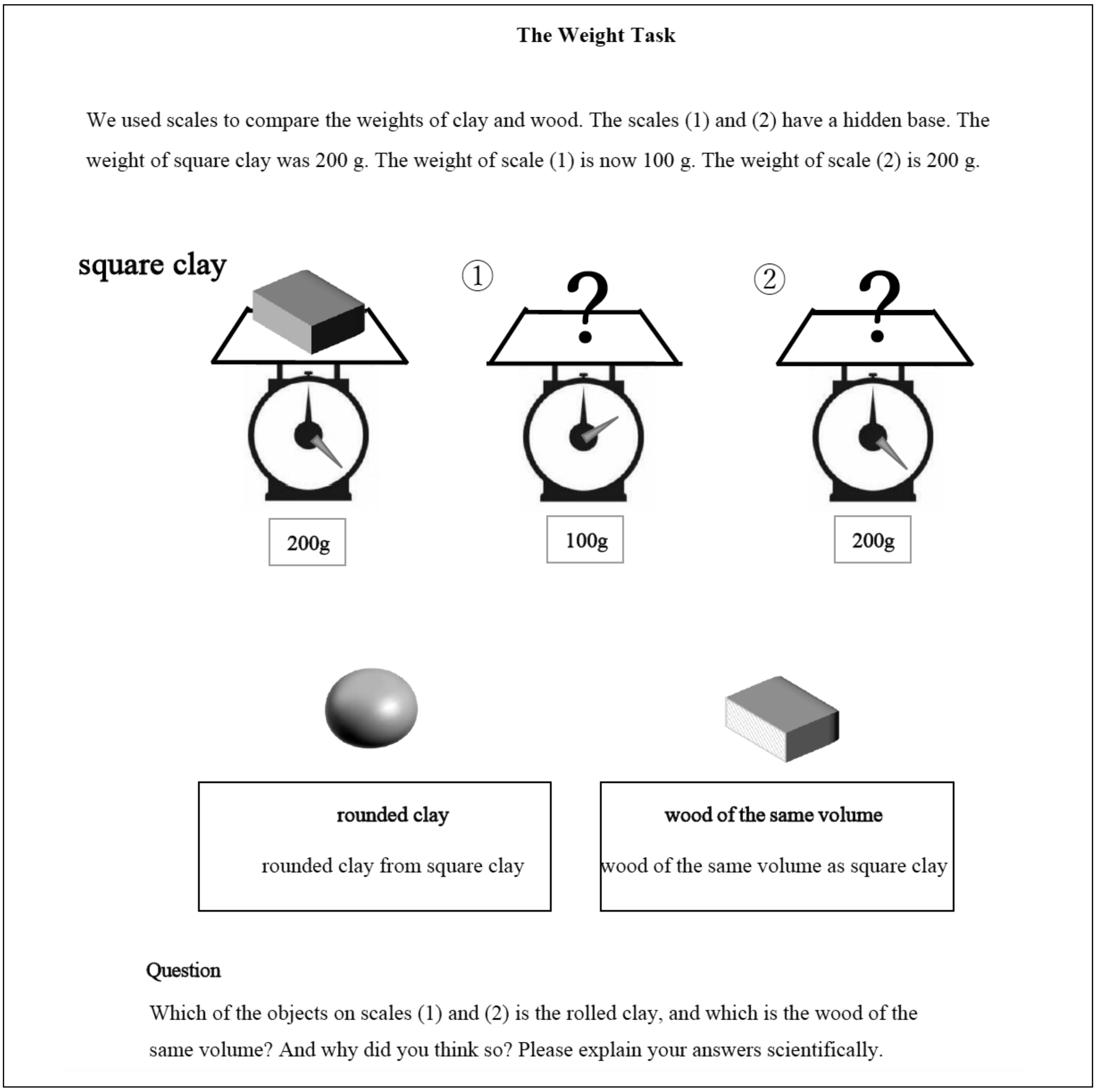



**2. The weight task**

Figure 6 illustrates the weight task, in which the following scales are presented: (1) a scale with a square of clay on it that points to a 200 g scale, (2) a scale with an object hidden that points to a 100 g scale, and (3) a scale with an object hidden that points to a 200 g scale. The three scales are illustrated in the figure. In addition, as options for the items placed on the scale in (1) and (2), 'rounded clay (rounded clay from square clay)' and 'wood of the same volume (wood of the same volume as square clay)' are illustrated. Then, the students were asked to describe why the items placed on scale (1) or scale (2) were 'rounded clay' or 'wood' of the same bodyweight or vice versa.


https://doi.org/10.33225/jbse/22.21.445

**Figure 6**
*The Weight Task*




*Data Analysis*

To examine improvements in the students' written arguments, each component was scored based on the rubric for the arguments that the students wrote freely before and after the class. First, the components of claim, evidence, and reasoning were checked, and if there was no description, the score was 0. In addition, for each of the three components, a score of 1 was given for an accurate description, and a score of 0 was given for a non-valid description. The following is a list of the descriptions that received 1 point for the ground task and the weight task. The scoring was done by two independent researchers (i.e., by one of the authors and one researcher specialising in science education), and an agreement rate of over 90% was confirmed.

1. **The ground task**
   Table 2 illustrates the rubric for the ground task. The arguments that received 1 point for claims were those that described either one or both of the following claims: 'The ground in A is in the sun', or 'The ground in B is in the shade'.


https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Arguments that were awarded one mark for evidence were those that described either one or both points of evidence for ground A and ground B, such as 'The temperature of the ground in A is 15°C at 9 am and 21°C at noon', or 'The temperature of the ground in B is 14°C at 9 am and 16°C at noon'. It was found that the observed data (temperature of ground A and temperature of ground B) was referred to as a claim, not evidence. Those that stated that 'the ground in A is hotter' and 'the ground in B is cooler' without mentioning the temperature, and those that stated that 'the ground in A is hotter than the ground in B' were also given 1 point because the arguments were based on the temperature.

The arguments that received 1 point for reasoning were those that included one or both of the following reasons for changes in ground temperature: 'Because sunlight heats the ground that is in the sun' and 'Because sunlight does not heat the ground in the shade'. It was found that the students' reasoning referred to temperature changes with respect to sun and shade. Therefore, statements such as 'In the sun, the sunlight falls on you' were scored 0 because they failed to refer to temperature changes.

The following is an example of a student's argument, in which 1 point was given for each component. In the argument, the statement 'Ground A is in the sun and ground B is in the shade' is a claim. The student can mention which ground is in the sun and which is in the shade. The statement 'The temperature of the ground in A is 15°C at 9:00 am and 21°C at noon, and the temperature of the ground in B is 14°C at 9:00 am and 16°C at noon, and in both cases, the temperature of ground A is higher than that of ground B' is evidence. The statement 'In the sun, the ground gets warm because of the sunlight, and in the shade, the ground does not get warm because there is a lack of sunlight' is the reasoning. The student can associate temperature change with the availability and non-availability of sunlight.

**2. The weight task**

Table 3 illustrates the rubric for the weight task. Arguments that received 1 point for a claim were those that described one or both of the following claims: 'What is placed on the scale in (1) is wood of the same volume', or 'What is placed on the scale in (2) is rounded clay.

The arguments that received 1 point for evidence were those that described either one or both scales in (1) and (2), such as 'The object on scale (1) weighs 100 g' and 'The object on scale (2) weighs 200 g'.

The arguments that received 1 point for reasoning were those in which either one or two of the reasons related to objects and weight, such as 'the weight of the objects does not change with the change in shape' or 'objects can have the same volume but different weights', were stated.

**Table 2**
*The Rubric for the Ground Task*

| Components | Points | Criteria |
|---|---|---|
| Claim | 1 | One or both of the following statements are true.<br>Ground A is in the sun.<br>Ground B is in the shade. |
| | 0 | No claims made. |
| Evidence | 1 | One or both of the following statements are made.<br>Example 1<br>The temperature of ground A is 15°C at 9 am and 21°C at noon.<br>The temperature of ground B is 14°C at 9 am and 16°C at noon.<br>Example 2<br>The temperature of ground A increased by 6°C.<br>The temperature of ground B increased by 2°C.<br>(If a student writes that 'the temperature change is large' or 'the temperature change is small' without indicating the numerical value, 1 point will be awarded).<br>Example 3<br>The temperature of ground A is higher than that of ground B. |
| | 0 | No evidence provided. |


https://doi.org/10.33225/jbse/22.21.445

| Components | Points | Criteria |
|---|---|---|
| Reasoning | 1 | One or both of the following statements are present.<br>Example 1<br>In the sun, the ground is warmed by sunlight.<br>In the shade, there is no sunlight, and the ground does not get warm.<br>Example 2<br>Because the temperature is higher in the sun.<br>Because the temperature is lower in the shade.<br>Example 3<br>Because the temperature is higher in the sun than in the shade. |
| | 0 | No reasoning provided. |

**Table 3**
*The Rubric for the Weight Task*

| Components | Points | Criteria |
|---|---|---|
| Claim | 1 | One or both of the following statements are true.<br>The object placed on the scale (1) is wood of the same volume as square clay.<br>The object placed on scale (2) is rounded clay. |
| | 0 | No claims made. |
| Evidence | 1 | Either or both of the following statements are present.<br>The scale in (1) weighs 100 g.<br>The scale in (2) weighs 200 g. |
| | 0 | No evidence provided. |
| Reasoning | 1 | One or both of the following statements were stated.<br>The weight of an object does not change even if the shape changes.<br>Objects having the same volume can have different weights. |
| | 0 | No reasoning statement was mentioned. |

The following is an example of a student's argument in which 1 point was given for each component. In the argument, the statements '(1) is wood of the same volume as square clay' and '(2) is rounded clay' are claims, and the student can mention which was wood of the same volume and rounded clay. The statements 'The square of clay weighs 200g, and the scale in (1) weighs 100g' and 'The square clay weighs 200g, and the scale in (2) also weighs 200g' are evidence. The student can compare the weight indicated in (1) and (2) and the weight of the square of clay. The statements, 'Different objects have different weights even if they have the same volume' and 'The weight of an object does not change even if it changes its shape' are reasons. The relationships between volume and weight and between shape and weight are mentioned.

*Analysis of Difficulties*

The arguments of the students who scored 0 on each component in the writing task and had difficulties even after guidance was provided on how to write arguments during the lesson were analysed. Arguments characterised by inadequacy or incompleteness in components (claim, evidence, and reasoning) were categorised as A type. Inappropriate cases, which included errors or inappropriate content, were categorised as type B. In addition to these, two other types of cases characterised by lack of clarity regarding the components also emerged. One was a case of confusion between evidence and reasoning, which was categorised as type C. In this type, the reasoning was not based on any scientific principle but was a fact based on the experimental result. Another type was a case of confusion between claim and evidence and was categorised as type D. In this case, instead of revealing the data (facts) of the experiment's results as evidence, the student states it with an interpretation and repeats the claim that answered the question. Four types of arguments were obtained, and their frequencies were determined. Two independent researchers (one of the authors and one researcher specialising in science education) categorised and determined the frequency, and an agreement rate of over 90% was confirmed.


https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

**Research Results**

*Analysis of Improvement in Written Arguments*

Sixty-seven of the target students were divided into two classes, and the differences in scores between students in the two classes were examined. Table 4 illustrates the pre-task distribution of the number of students by a score between the two classes. The claims, evidence, and reasoning of the two types of writing tasks were compared between classes 1 and 2 using Mann-Whitney's U test. As there were no significant differences between all components of the two tasks (*ns*), it was therefore decided that there would be no difference between the two classes, and an analysis of the writing task would be carried out on 67 students at the same time.

Table 5 illustrates the distribution of the number of participants by the score in the writing task. The scores in the three components of claim, evidence, and reasoning were compared between pre- and post-tasks.

**Table 4**
*The Pre-task Distribution of the Number of Students by the Score between the Two Classes*

| Task | Component | Class1 | | Class2 | | Z- value |
|---|---|---|---|---|---|---|
| | | 1 point | 0 point | 1 point | 0 point | |
| Ground | Claim | 29 | 4 | 30 | 4 | .045 |
| | Evidence | 20 | 13 | 19 | 15 | .389 |
| | Reasoning | 25 | 8 | 29 | 5 | .979 |
| Weight | Claim | 30 | 3 | 25 | 9 | 1.841 |
| | Evidence | 3 | 30 | 2 | 32 | .496 |
| | Reasoning | 25 | 8 | 23 | 11 | .731 |

**Table 5**
*Distribution of the Number of Participants by the Score in the Writing Task*

| Task | Component | Pre-task | | Post-task | |
|---|---|---|---|---|---|
| | | 1 point | 0 point | 1 point | 0 point |
| Ground | Claim* | 59 | 8 | 66 | 1 |
| | Evidence** | 39 | 28 | 66 | 1 |
| | Reasoning | 54 | 13 | 56 | 11 |
| Weight | Claim** | 55 | 12 | 66 | 1 |
| | Evidence** | 5 | 62 | 49 | 18 |
| | Reasoning** | 48 | 19 | 63 | 4 |

**$p < .01$; *$p < .05$

1. **The ground task**
   First, the pre-test and post-test scores for each component (claim, reasoning, and evidence) were compared. When scores on the pre-test and post-test for each component were compared using the McNemar test, a significant improvement in scores was observed for claim and evidence (claim, $p < .05$, $M$=4.000; evidence, $p < .01$, $M$=25.037). However, there was no significant improvement in the reasoning (*ns*, $M$=0.071), although more than 80% of the students obtained a perfect score in the post-test task. On comparing the distribution of the number of students for each component in the post-task, it was found that the number of students who got 1 point for claim and evidence was high, while the number of students who got 1 point for reasoning was low.


https://doi.org/10.33225/jbse/22.21.445

2. **The weight task**

First, the pre-test and post-test scores on each component were compared. When the scores for each of the components were compared between the pre-test and post-test tasks using the McNemar test, a significant improvement in scores for all the three components (claim, $p < .01$, $M$=7.692; evidence, $p < .01$, $M$=40.196; reasoning, $p < .01$, $M$=9.333) was found. On comparing the distribution of the number of students for each component in the post-task, the number of students who got 1 point for claim and reasoning was high, while the number of students who got 1 point for evidence was low.

*Analysis of Difficulties*

In the analysis of the writing task, a total of 27 students who scored 0 on at least one of the components of the ground task and the weight task on the post-test were analysed for their arguments. It was found that the arguments that got a score of 0 in the post-test task could be categorised into four types of arguments that had some inadequacies in common between the ground task and the weight task. These four types of arguments (A–D), based on the argument structure and distribution of the number of participants for each type, have been depicted below in Table 6. In the next section, the details of each type based on the arguments in the post-task are discussed.

**Table 6**
*Distribution of the Number of People by the Type of Difficulties in the Post-task*

| Task | Type A | Type B | Type C | Type D |
|---|---|---|---|---|
| | Incompleteness of components | Inappropriateness of components | Confusion between evidence and reasoning | Confusion between claim and evidence |
| Ground | 1 | 4 | 8 | 0 |
| Weight | 2 | 4 | 0 | 17 |

1. **Type A: Incompleteness of components**

Type A included arguments that lacked any of the claim, evidence, or reasoning components. In the ground task, there was one case that lacked reasoning. In the weight task, one such argument was found in the claim category and one case in evidence. Specifically, in the ground task, the statements that lacked reasoning were: 'The temperature of ground A is 15°C at 9 am and 21°C at noon, and the temperature of ground A is higher at noon than at 9 am because the sun is directly overhead during noon. The temperature of ground A is high because the sun is directly overhead at noon compared to 9 am. The temperature of ground in B is low because it is 14°C at 9 am and 16°C at noon. Therefore, ground A is in the sun and ground B is in the shade'. The argument and evidence were provided, but the reasoning was missing.

2. **Type B: Inappropriateness of component**

Type B included arguments that contained errors or inappropriate content. In the ground task, one such case was found regarding claims, one for evidence, and two cases for reasoning, while in the weight task, four cases were found pertaining to the reasoning component. In the ground task, regarding the claim, there were statements such as 'ground A is in the shade' and 'ground B is in the sun', and the correspondence between the ground in A and the ground in B (and between the sun and the shade) was incorrect. With respect to the evidence, there were incorrect statements that did not focus on the temperature of the ground, such as 'there is more light at noon'. With respect to the reasoning, the answer was limited to 'because, in the sun, the sun hits you', and no mention of temperature change was made. In the weight task, answers such as 'wood is lighter', and 'wood and clay are different things' lacked the necessary description about comparing the weight of objects having the same volume.

3. **Type C: Confusion between evidence and reasoning**

Type C refers to arguments that indicated students' confusion regarding evidence and reasoning. In these cases, the provided reasoning lacked a scientific basis but was based on the observed experimental results. Eight such cases were identified in the ground task that pertained to the reasoning component. The following is an example of a type C argument. Specifically, the statement, 'Why can we say the answer

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

with this evidence, because…' confuses evidence with reasoning, describing what constitutes evidence, such as 'the temperature of the ground in A is higher', even though it is trying to explain the reasoning.

> Ground A is in the sun, and ground B is in the shade. The evidence is that the temperature of ground A at noon is 20°C and ground B is 16°C. Why can we provide this answer with this evidence? <u>Because the temperature of the ground in A is higher.</u>

4. **Type D: Confusion between claim and evidence**

Type D refers to arguments that confused claims with evidence. In evidence statements of the weight task, 17 such cases were identified. The following is an example of a type D argument. Although the student is trying to explain the evidence with the statement 'Evidence is…', they do not state the experimental result (fact) that 'The weight of the object placed on scale (2) is 200 g'. Instead, the student states the interpretation as follows: 'the rounded clay weighs 200 g' and repeats the claim as to which is the rounded clay.

> (1) has the same volume as wood, and (2) is the rounded clay. The evidence is that <u>the square clay weighs 200 g and the rounded clay also weighs 200 g.</u> This can be used as evidence to explain that the weight does not change with the change in the shape of an object.

## Discussion

### *Improvement and Difficulties of Students' Written Arguments*

The first research question is: 'Can science lessons based on Zembal-Saul et al.'s writing scaffold improve the written arguments of third-grade primary school students (aged 9–10)'. Comparing the scores obtained on the pre-test and post-test writing tasks, the number of students who scored 1 point in each component increased, and there was a significant improvement in the scores on almost all the components. Therefore, it is clear that third-grade students can improve their written arguments through argument teaching.

The second research question is: 'What types of students' difficulties emerge after learning argumentation'. Some difficulties were revealed through the students' arguments. The first was the lack of certain components that form an argument (type A). This is consistent with the difficulties pointed out by McNeill and Krajcik (2011). Yamamoto et al. (2013) pointed out that the students in the upper grades of primary school (11-years old) also tend to omit evidence, considering it obvious from the result. The second type of difficulty concerns errors or inappropriate content (type B), as McNeill and Krajcik (2011) pointed out. Instances of type B depict a lack of understanding of the content, misinterpretation of claims and evidence, and reasoning based on everyday life experiences instead of scientific principles. The third type of difficulty is the confusion between evidence and reasoning (type C), which was evident in the ground task. The reason might be that the relationship between sunlight and heat (taken as a scientific principle) was more difficult for students to understand as compared to the conservation of mass. The fourth type of difficulty concerns the confusion between claim and evidence (type D), which was evident in the weight task. It may have been more difficult for the students to use the event (numerical values illustrated by the scales) as evidence than the graphical data. Although the case of McNeill and Krajcik (2011) does not reveal such confusion, Berland and Reiser (2009) have pointed out that even middle school students have difficulty distinguishing between assertions and evidence or reasoning, which was also evident in the 9-year-old students in this study. In particular, the confusion between evidence and reasoning was also pointed out in the survey questions of Trends in International Mathematics and Science Study (TIMSS) 2003 and overlaps with the problem of learners' difficulties in distinguishing between facts and reasons.

### *Suggestions for Teaching to Overcome Difficulties*

The data obtained from this study may be insufficient to determine how to overcome the various types of difficulties students face, which may depend on the context of the task. However, as mentioned above, the difficulties found can be interpreted in the light of previous research and recognised as more general difficulties, and it is important to recognise the potential of further teaching to help alleviate these difficulties.

To overcome the first difficulty (incompleteness of components), it is necessary to make the students aware of the necessity of explaining the components to others, without omitting them, during free writing when the scaffolding provided begins to fade. At the same time, it is also necessary to provide the students with detailed


https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/


teaching repeatedly so that they can verify the presence and accuracy of the components. The fading of the scaffolding from the sheet, which requires students to describe each component in a frame, to free writing will pose a great challenge to the students. It is difficult to keep track of all the components during free writing, and as a result, it can be inferred that students compose incomplete arguments without being aware of the lack of certain components. Zembal-Saul et al. (2012) made the students state scientific explanations in their notebooks through free writing without dividing them into different components. The students were asked to point out which part of the explanation was the claim, evidence, or reasoning using a post-it and asked to revise their notes to include any explanation they might have missed (p. 108). While incorporating the teaching methods, it is necessary to repeat the teaching that encourages self-evaluation and mutual evaluation by paying attention to the presence and correctness of each component by, for example, having the students underline the parts of their claims in free writing, using dotted lines for evidence and wavy lines for reasoning.

The second difficulty (incorrect or inappropriate content about the components) was closely related to understanding the content, demonstrating that more attention needs to be paid to 'content support'. McNeill (2011) noted that 'If students do not understand the science concepts, they are unable to construct a scientifically accurate claim or appropriately justify that claim' (p. 820). One of the purposes of argumentation is to understand scientific concepts (Berland & Reiser, 2009), and since understanding leads to adequate statements of the argument, the two have a complementary relationship. While teaching argumentation in the classroom, it is necessary to provide relevant content support for students to better understand the claims that are justified by the experiment.

Finally, with respect to the third difficulty (confusion between evidence and reasoning) and the fourth difficulty (confusion between claim and evidence), guidance was needed to clarify the distinction between the components in 'general and content support' and 'structure'. The confusion between evidence and reasoning was possibly due to the inability of the students to distinguish between facts and interpretations. McNeill and Krajcik (2011) encouraged the use of scientific principles in reasoning, but in the experimental activity, the scientific principle for the student is the hypothesis that is formulated that depicts the cause of the event. If the experiment results support this hypothesis, it becomes a scientific principle for reasoning. It is emphasised here that the hypothesis that is supported by the experiment is a scientific principle. In addition, the confusion between the claim and the evidence may occur when students are not aware that evidence is a 'fact' apparent in the result of an experiment, and hence they may respond by adding an interpretation. To facilitate students' understanding of the evidence present in an argument, it is necessary to make them aware that while stating the evidence, the numbers and results presented as data should be presented as they are, without omitting them as facts.

## Conclusions and Implications

This study demonstrated that it is possible to improve the written arguments of third-year primary school students (aged 9–10). The finding that even lower primary school students can construct an argument consisting of claims, evidence, and reasoning with appropriate teaching may encourage the promotion of written argument teaching in primary schools from an early stage.

Another implication of this study is that four difficulties remained for students even after the teaching, which can be described as 'strong difficulties'. Since the students' difficulties found in this study were similar to those reported in previous studies of upper primary school students (fifth-grade and above) and middle school students, common teaching methods are needed to overcome these difficulties in teaching written arguments. Therefore, it is important to focus on teaching strategies to overcome the four difficulties. Future teaching strategies focusing on the types of difficulties identified in this study can be used to improve students' written arguments.

As student difficulties may depend on the context of the tasks, data from a small, limited-content study such as this are limited. Further research is needed to develop teaching strategies that encompass the different types of difficulties faced by students. In the future, it will be necessary to develop such teaching strategies for other units of teaching and analyse their effects on students' ability to construct written arguments through repeated teaching.

## Declaration of Interest

Authors declare no competing interest.


https://doi.org/10.33225/jbse/22.21.445

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

**Acknowledgements**

The authors would like to express their deepest gratitude to the principal, students, and parents of Elementary School Attached to Kobe University for their cooperation in conducting this research. This study was supported by JSPS KAKENHI Grant Number JP18K02936.

**References**

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26–55. https://doi.org/10.1002/sce.20286

Chin, C., & Osborne, J. (2010). Supporting argumentation through students' questions: Case studies in science classrooms. *The Journal of the Learning Sciences*, *19*(2), 230–284. https://doi.org/10.1080/10508400903530036

Clark, D. B., & Sampson, V. D. (2007). Personally-seeded discussions to scaffold online argumentation. *International Journal of Science Education*, *29*(3), 253–277. https://doi.org/10.1080/09500690600560944

Çoban, Ü. G. (2013). The effects of inquiry supported by argument maps on science process skills and epistemological views of prospective science teachers. *Journal of Baltic Science Education*, *12*(3). 271–288. https://doi.org/10.33225/jbse/13.12.271

Erduran, S., & Jiménez-Aleixandre, M. P. (Eds.). (2007). *Argumentation in science education: Perspectives from classroom-based research*. Springer.

Hong, Z.R., Lin, H.S., Wang, H.H., Chen, H.T. & Yang, K.K. (2013). Promoting and scaffolding elementary school students' attitudes toward science and argumentation through a science and society intervention. *International Journal of Science Education*, *35*(10), 1625–1648. https://doi.org/10.1080/09500693.2012.734935

Iordanou, K. (2010). Developing argument skills across scientific and social domains. *Journal of Cognition and Development*, *11*(3), 293–327. https://doi.org/10.1080/15248372.2010.485335

Iordanou, K., & Constatinou, P. (2015). Supporting use of evidence in argumentation through practice in argumentation and reflection in the context of SOCRATES learning environment. *Science Education*, *99*(2), 282–311. https://doi.org/10.1002/sce.21152

Iordanou, K., Kuhn, D., Matos, F., Shi, Y., & Hemberger, L. (2019). Learning by arguing. *Learning and Instruction*, *63*, Article Number: 101207. https://doi.org/10.1016/j.learninstruc.2019.05.004

Kamiyama, S., Yamamoto, T., Yamaguchi, E., Sakamoto, S., Muratsu, K., & Inagaki, S. (2015). Instructional strategies for teaching primary students to construct arguments with rebuttals. In J. Lavonen, K. Juuti, J. Lampiselkä, A. Uitto & K. Hahl (Eds.), *Electronic Proceedings of the ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future, Part 7* (co-ed. Maria Andrée & Maria Pilar Jimenez-Aleixandre), (pp. 997–1003).

Kaya, E. (2013). Argumentation practices in classroom: Pre-service teachers' conceptual understanding of chemical equilibrium. *International Journal of Science Education*, *35*(7), 1139–1158. https://doi.org/10.1080/09500693.2013.770935

Kuhn, D. (2010). Teaching & learning science as argument. *Science Education*, *94*(5), 810–824. https://doi.org/10.1002/sce.20395

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, *22*(4), 545-552. https://doi.org/10.1177/0956797611402512

Kuhn, D., Hemberger, L., & Khait, V. (2016). Tracing the development of argumentive writing in a discourse-rich context. *Written Communication, 33*(1), 92–121. https://doi.org/10.1177%2F0741088315617157

McNeill, K. L. (2009). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education*, *93*(2), 223–268. https://doi.org/10.1002/sce.20294

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching*, *48*(7), 793–823. https://doi.org/10.1002/tea.20430

McNeill, K. L., & Krajcik, J. (2011). *Supporting grade 5–8 students in constructing explanations in science.* Pearson.

Monteira, S.F., & Jiménez-Aleixandre, M.P. (2016). The practice of using evidence in kindergarten: The role of purposeful observation. *Journal of Research in Science Teaching*, *53*(8), 1232–1258. https://doi.org/10.1002/tea.21259

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, *41*(10), 994–1020. https://doi.org/10.1002/tea.20035

Osborne, J. F., Henderson, J.B., MacPherson, A., Szu, E., Wild, A., & Yao, S.Y. (2016). The development and validation of a learning progression for argumentation in science. *Journal of Research in Science Teaching*, *53*(6), 821–846. https://doi.org/10.1002/tea.21316

Ping, I. L. L., Halim, L., & Osman, K. (2020). Explicit teaching of scientific argumentation as an approach in developing argumentation skills, Science process skills and biology understanding. *Journal of Baltic Science Education*, *19*(2), 276–288. https://doi.org/10.33225/jbse/20.19.276

Pontecorvo, C., & Girardet, H. (1993). Arguing and reasoning in understanding historical topics. *Cognition and Instruction*, *11*(3–4), 365–395. https://doi.org/10.1080/07370008.1993.9649030

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, *88*(3), 345–372. https://doi.org/10.1002/sce.10130

Saracaloglu, A.S., Aktamis, H., & Delioglu, Y. (2011). The impact of the development of prospective teachers' critical thinking skills on scientific argumentation training and on their ability to construct an argument. *Journal of Baltic Science Education*, *10*(4), 243–260. http://oaji.net/articles/2014/987-1410547733.pdf

Toulmin, S. (1958). *The use of argument.* Cambridge University Press.


https://doi.org/10.33225/jbse/22.21.445

Yamamoto, T., Sakamoto, M., Yamamguchi, E., Nishigaki, J., Muratsu, K., Inagaki, S., & Kamiyama, S. (2013). Teaching strategies of argument to elementary school children: Through practice in the unit of 'pendulum movement' (in Japanese). *Journal of Research in Science Education*, *53*(3), 471–284. https://doi.org/10.11639/sjst.KJ00008877262
Zembal-Saul, C. L., McNeill, K. L., & Hershberger, K. (2012). *What's your evidence?: Engaging K-5 children in constructing explanations in science*. Pearson.

Received: *March 16, 2022* Revised: *April 18, 2022* Accepted: *May 06, 2022*

Cite as: Yamamoto, T., Kamiyama, S., Tanaka, T., & Yamaguchi, E. (2022). Primary school students' difficulties in writing arguments: Identifying challenges and opportunities for science teaching. *Journal of Baltic Science Education*, *21*(3), 445-461. https://doi.org/10.33225/jbse/22.21.445

| | |
|---|---|
| ***Tomokazu Yamamoto*** *(Corresponding author)* | PhD, Professor at the Hyogo University of Teacher Education, 942-1 Shimokume, Kato, Hyogo, 6731494, Japan. E-mail: ttyamamo@hyogo-u.ac.jp Website: https://hute-rd.hyogo-u.ac.jp/profile/en.7561f4d63aabcaf7028c7b24fbed6cd7.html ORCID: https://orcid.org/0000-0002-2300-3736 |
| ***Shinichi Kamiyama*** | PhD, Associate Professor at the Osaka University of Health and Sport Sciences, Asashirodai, Kumatori Sennan, Osaka, 5900496, Japan. E-mail: kamiyamashinichi24@gmail.com Website: https://www.ouhs.jp/department/teacher/kamiyama_s/ ORCID: https://orcid.org/0000-0002-6676-0438 |
| ***Tatsuya Tanaka*** | A teacher at the Kobe Municipal Nadanohama Primary School, 2-2-1, Mayakaigandori Nada, Kobe, 6570855, Japan. E-mail: tanakatatsuya1204@gmail.com |
| ***Etsuji Yamaguchi*** | PhD, Professor at the Graduate School of Human Development and Environment, Kobe University, 3-11 Tsurukabuto Nada, Kobe, 6578501, Japan. E-mail: etuji@opal.kobe-u.ac.jp Website: https://researchmap.jp/etsujiyamaguchi?lang=en ORCID: https://orcid.org/0000-0001-5140-1136 |