International Electronic Journal of Mathematics Education
2025, 20(3), em0830
e-ISSN: 1306-3030
https://www.iejme.com

MODESTUM

Research Article **OPEN ACCESS**

Enhancing critical thinking in mathematics education: A rubric for students' social values

Yuichiro Hattori [1*], Yuuki Inoue [2], Kazuki Matsubara [3], Ryoto Hakamata [4], Yoichiro Hisadomi [5]

[1] Okayama University, Okayama, JAPAN
[2] Hiroshima University High School, Hiroshima, JAPAN
[3] Saitama University, Saitama, JAPAN
[4] Kochi University, Kochi, JAPAN
[5] Hiroshima Prefectural Education Center, Hiroshima, JAPAN
***Corresponding Author:** hattori.yuichiro@okayama-u.ac.jp

Citation: Hattori, Y., Inoue, Y., Matsubara, K., Hakamata, R., & Hisadomi, Y. (2025). Enhancing critical thinking in mathematics education: A rubric for students' social values. *International Electronic Journal of Mathematics Education, 20*(3), em0830. https://doi.org/10.29333/iejme/16186

ARTICLE INFO

Received: 30 May 2023
Accepted: 02 Mar 2025

ABSTRACT

This research focuses on enhancing critical thinking in secondary mathematics education using carefully constructed teaching materials. Utilizing socially open-ended framework, complemented by perspectives from critical mathematics education, the study aims to promote students' critical thinking skills. The developed materials were implemented in a unique lesson centred around the 60-second challenge, whereby students were tasked to propose rule modifications. The study also aimed to identify students' social values, considering these as interconnected with their capacity for critical thought. The assessments involved multiple teachers evaluating student reports using rubric and students participating in peer review activities within the class. The lesson's outcome illustrated students' ability to propose thoughtful improvements to the existing rules, reflecting their critical thinking abilities. Furthermore, this innovative series of evaluation activities via the rubric provides a foundation for future research in the assessment of critical thinking, serving as a potential model for future studies.

Keywords: secondary mathematics education, critical thinking, critical mathematics education, the socially open-ended problem, rubric

INTRODUCTION

Research on critical thinking, a universal skill, has been extensively conducted in various fields such as education, psychology, and philosophy (Ennis, 1987; Paul, 1992). However, there is still a lack of pedagogical research specifically focused on mathematics education, both globally and in Japan (Baba, 2016, 2017, 2019, 2020; Fonseca & Arezes, 2017; Jablonka, 2020). In Japan, the emphasis of mathematics education research has primarily been on developing teaching materials and instructional methods to foster students' critical thinking (Hattori & Fukuda, 2019; Kubo, 2019). However, new challenges have emerged, such as properly identifying and evaluating students' critical thinking abilities. This research question remains largely unexplored and urgent, not only within the academic education field in Japan but also internationally.

This study adopts the perspective of critical thinking in the context of critical mathematics education and establishes connections with the societal values expressed by individuals. As Skovsmose (2020) proposed, critical mathematics education emphasises the importance of addressing oppression, exclusion, and exploitation in society through educational approaches. It is characterised by its pursuit of social justice, commitment to opening new possibilities for students, and a critical approach to mathematics in confronting all forms of oppression. Furthermore, critical mathematics education encompasses a broad perspective that considers not only explicit issues but also education for all students, making it a concept that transcends narrow applications (Skovsmose, 2016, 2022).

The 'spirit of harmony' (*wa no seishin*), a hallmark of Japanese culture, can sometimes lead to peer pressure wherein individuals with minority opinions are implicitly pressured to align with the majority. This tendency to suppress personal opinions is a characteristic of Japanese society; however, various forms of oppression are interconnected (Benedict, 1946). Moreover, oppression manifests differently across societies (Kincheloe et al., 2017). In Japan, peer pressure represents a pervasive form of oppression that is subtly and widely embedded in the fabric of society. Overcoming this form of oppression without compromising harmony or kindness is a critical challenge. When individuals feel that their ideas are respected, they are more likely to engage in society and maintain motivation to learn (Frankenstein, 2012).

Table 1. Comparison of the two types of open-ended problems (Baba, 2007, p. 22)

	Mathematically open-ended problem	Socially open-ended problem
Objective	To nurture mathematical thinking	To nurture mathematical thinking and judgement based on mathematical thinking and associated social values
Problem	To allow mathematically diverse solutions	To allow mathematically diverse solutions and associated social values
Method	Discussion on mathematically diverse solutions and their generalisation and symbolisation	Discussion on mathematically diverse solutions and the associated social values

In this sense, overcoming 'peer pressure' is the key to achieving critical mathematics education in Japan. As a methodological approach, this study adopted the framework of socially open-ended problems. As Baba (2009) proposed, socially open-ended problems encourage students to articulate their social values and engage in mathematical problem solving based on these values. This approach fosters a classroom culture in which even small, yet meaningful expressions and problem-solving efforts are celebrated, reflecting the unique characteristics of Japanese educational practices. Additionally, we have created teaching materials with the purpose of fostering critical thinking and have assessed their effectiveness through practical implementation. To evaluate their effectiveness, we will utilise a rubric as a measurement tool. Engaging in socially open-ended problems with critical thinking requires learners to be aware of their own societal values and engage in reasoning and discussions concerning the societal values of others. In other words, social values must be somehow manifested in problem-solving endeavours; otherwise, awareness, sharing, reasoning, and discussions would be rendered impossible. Thus, the "manifestation of social values" becomes a crucial factor. Establishing a rubric will be necessary to assess this aspect. During the lessons, multiple teachers evaluated student reports using the rubric, and students themselves participated in peer evaluation activities within the classroom. We anticipate that this series of evaluation activities, utilising the rubric, will provide fundamental data for research on evaluating the development of critical thinking in mathematics education. The objectives of this study, therefore, are to develop teaching materials that promote critical thinking in secondary mathematics education, describe students' problem-solving approaches through practical implementation, and test the effectiveness of a rubric in capturing the societal values expressed by students.

THEORETICAL FRAMEWORK

Critical Mathematics Education and Socially Open-Ended Problem

We have adopted Skovsmose's (1994) perspective on critical mathematics education. According to this perspective, educational practices within critical mathematics education aim to cultivate critical citizenship within the context of society. In critical mathematics education, mathematics serves as both a tool for critique and an object of criticism within the social context. It is positioned as a valuable means to identify and analyse significant aspects of society (Skovsmose & Nielsen, 1996). The teaching and learning process should strive to equip students with critical capabilities that are essential for their participation in the ongoing democratisation of society (Skovsmose, 1994). Critical mathematics education can be characterised by its pursuit of social justice, its presentation of new possibilities to students, and its critical engagement with mathematics in various forms and applications (Skovsmose, 2020). Essentially, it is a comprehensive concept that must be regarded as education for all students, rather than merely a tool for critiquing systems of social inequality and oppression (Skovsmose, 2016, 2022). In this study, we consider the school environment, where students typically reside, as their own society, and our aim is to critically examine the use of mathematics to enhance the school community. Skovsmose (1994) describes the purpose of pedagogical practices within critical mathematics education as follows:

> …not to concretise mathematics but to see in what way mathematics could develop from a broad context requiring the use of mathematics. If mathematics is everywhere in daily life situations, then it need not be necessary to develop any artificial concretisation. Instead, the prerequisite must be to create open-ended situations and in these, to let mathematics grow (p. 79).

To achieve this objective and create a harmonious framework, we employ socially open-ended problems (Baba, 2007; Baba & Shimada, 2019). Socially open-ended problems are defined as problems that have multiple mathematical and social solutions, with the aim of developing social judgment through mathematical approaches (Baba, 2009, p. 52). **Table 1** illustrates the distinctions between socially open-ended problems and mathematically open-ended problems. As depicted in **Table 1**, mathematically open-ended problems focus on the development of mathematical thinking itself, whereas socially open-ended problems prioritise the development of social judgment, utilising mathematics solely as a problem-solving tool.

Research on socially open-ended problems has recently gained attention, with studies being conducted among high school students in Thailand that highlight its growing prominence and the potential for further accumulation of findings (González & Chitmun, 2017). Recent research has focused on ethics in relation to socially open-ended problems (Hattori et al., 2021).

In this study, it is essential to clarify the meaning of values. Values are the guiding principles that shape human behaviour (Sagiv et al., 2017). Socially open-ended problems entail the utilisation of diverse mathematical models that are built upon the children's values inherent in the problem-solving process. Consequently, by encouraging students to recognise, appreciate, and critically examine these values, we can foster the development of their critical thinking skills.

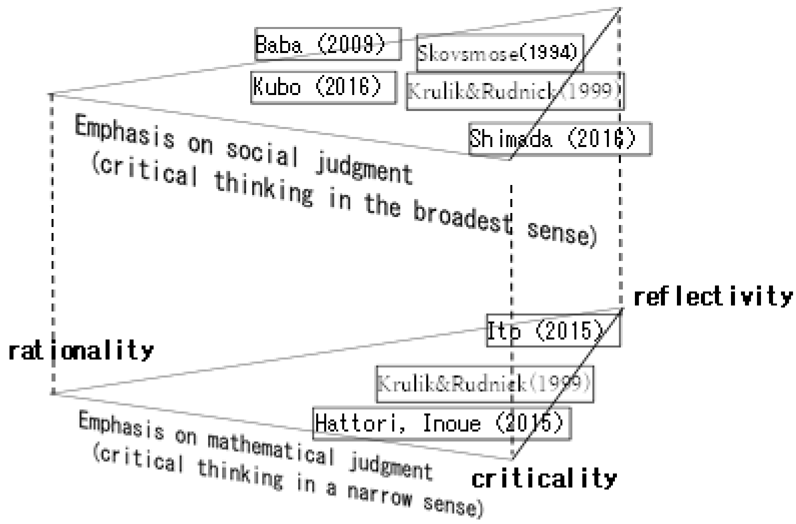


Figure 1. Literature review on the development of critical thinking in mathematics education (Source: Hattori, 2017; translated into English)

Critical Thinking

Critical thinking is vital for 21st-century individuals (Vincent-Lancrin et al., 2019). However, its definition is multifaceted across different research fields, lacking a fixed definition (Michita, 2003). We adopt Paul's (1995) model for this study, focusing on 'fairness', 'morality', and 'society'. Paul distinguishes 'strong' and 'weak' critical thinking. 'Strong' thinkers, according to Paul (1992), have seven traits: humility, courage, empathy, good faith, perseverance, faith in reason, and a sense of justice. Michita (2005) characterised 'strong' critical thinkers as recognising empathetic, fair evaluators who value different viewpoints, even when they challenge their beliefs. Conversely, 'weak' critical thinkers identify and refute others' flaws, supporting their own beliefs (Paul, 1995). In this study, we aim to cultivate 'strong' critical thinking, enabling students to engage with others' opinions constructively and empathetically, while critically reflecting on their own ideas. 'Strong' critical thinking enables fair judgments through consideration of diverse interests (Michita, 2005).

Next, we explore the notion of critical thinking within the mathematics education context. Critical thinking plays various roles in mathematics education, including viewing it as a byproduct of mathematics learning, prerequisite for mathematical problem-solving, and means of critically engaging with social, political, and environmental issues (Jablonka, 2020). Geiger et al. (2023) highlighted the significance of critical mathematical thinking (CMT) in mathematics education, emphasising its importance beyond skill acquisition. They defined CMT as the ability to critically analyse real-world problems and make effective and ethical decisions. Furthermore, understanding how critical thinking connects with values when addressing the role of mathematics in society is an essential research topic (Kacerja & Julie, 2023). This study approaches critical thinking from the perspective of solving real-world mathematical problems. As a methodological aspect of fostering such critical thinking, this study focuses on socially open-ended problems that emphasise student values.

Michita (2013) identified critical thinking through three key concepts: rationality, reflectivity, and criticality. Different emphasis on each aspect can alter interpretations. Hattori (2017) applied this understanding to highlight the importance of critical thinking in mathematics education from two perspectives: critical thinking in a broader sense and critical thinking within the pure mathematical context (referred to as critical thinking in a narrower sense). Furthermore, Hattori (2017) organised the existing studies on fostering critical thinking in mathematics education (**Figure 1**), indicating that social values are expressed in the classroom during critical thinking in the broader sense, while mathematical values emerge during critical thinking in the narrower sense.

The definitions of critical thinking in the broad and narrow senses are provisionally outlined as follows.

Critical thinking in the broadest sense

In this study, we tentatively adopt Hattori's (2020) definition of critical thinking in the broad sense:

> Students should be able to see the essence of a given problem without being misled by superficial differences, and to propose a solution based on a mathematical model constructed from their own values. Also, while judging the validity and reliability of solutions proposed by others, students should try to modify their own solution proposal to improve it (Hattori, 2020, p. 162).

Education for critical thinking aims to overcome potentially limiting perspectives, thus necessitating the transcendence of the very criteria that define 'critical' judgment (Jablonka, 2020). This is why we employ the term 'tentatively'.

Critical thinking in the narrow sense

Using mathematical knowledge and reasoning to assess a situation's authenticity is crucial (Hattori & Inoue, 2015).

In classrooms, students use critical thinking and social values to solve open-ended problems. Students utilise critical thinking in conjunction with their own individual or group social values to seek solutions. This study aligns critical thinking with the broader interpretation depicted in **Figure 1**, aiming to enhance students' critical thinking by focusing on activities involving critical examination of others' values.

Developing the Socially Open-Ended Problem: Proposing New Rules for the 60-Second Challenge

Lesson overview and targeted critical thinking skills

This study focused on secondary mathematics education and aimed to create teaching materials that specifically targeted the development and evaluation of critical thinking. The teaching materials in question were designed as socially open-ended problems. More specifically, we developed a socially open-ended problem centred around the task of proposing new rules for the 60-second challenge. This problem was presented in a classroom setting with Japanese second-year high school students.

During the lesson, the students were tasked with examining the validity of the existing rules of the 60-second challenge, a recreational game typically played by elementary school students. Based on their evaluation, the students were required to propose new rules for the game. The teaching materials were developed as three-hour units, and the objective was to nurture specific critical thinking skills in the students.

The critical thinking skills we aimed to foster in the classroom were as follows: students should be capable of 'formulating reasonable rules for the 60-second challenge that can be widely accepted, while making their own values evident through a mathematical model', and 'critically evaluating the rules proposed by others and modifying their own rules to improve them'. These skills were tentatively adopted for the study.

These skills are grounded in the components of criticality and rationality, as illustrated in **Figure 1**. Criticality refers to the endeavour to create better rules using mathematics, whereas rationality signifies the creation of rules that everyone can agree upon. Furthermore, this study strongly emphasises critical thinking, as highlighted by Paul (1992). Specifically, the element of 'otherness' is a key aspect of evaluation. The ability to empathically understand others' perspectives is essential.

Students were expected to utilise their existing mathematical knowledge, primarily in the field of statistics, to analyse the nature of the data and establish new rules for the 60-second challenge. Additionally, they were tasked with addressing questions such as 'should the degree of variability be considered in determining the outcome of the game? (If so, how should it be incorporated?)' and 'should there be a handicap system based on grade level? (If so, how should it be implemented?)'. These questions allowed for diverse solutions based on each student's individual values.

Here, we will explain how examining the validity of the '60-Second Challenge' rules connects to critical thinking. As the chapter on the theoretical framework indicates, this study draws on critical mathematics education (Skovsmose, 1994) and aims to foster students' critical thinking skills. From the perspective of critical mathematics education, the goal is to use mathematics to address issues and contribute to building a better society. Evaluating the validity of the '60-Second Challenge' rules is linked to enriching the social environment of elementary schools where students are enrolled. In this sense, critical thinking that considers and empathises with others is expected, particularly among elementary school students. This aligns with what Paul (1992) referred to as critical thinking in its strong sense, emphasising the ability to empathise with others' opinions and constructively develop one's own ideas.

In this practice, the primary focus is on fostering critical thinking in a broad sense. However, proposing more sophisticated mathematical approaches signifies a narrow application of critical thinking, which refers to thinking within the domain of mathematics, specifically aimed at developing better mathematical practices. For instance, simply comparing the number of participants who stopped the stopwatch exactly at the 60-second mark is an inappropriate measure because it does not account for differences in group sizes. Instead, analysing the data using proportions calculated by dividing them by the number of participants exemplifies the application of critical thinking in a narrow sense, utilising more advanced mathematics.

Next, we will discuss the tasks that were deemed central by the students during the lesson.

Central tasks of the class

The 60-second challenge is a game in which participants close their eyes and start a stopwatch, attempting to stop it when they believe 60 seconds have elapsed. The participant with the time closest to 60 seconds is declared the winner. This game has been established as an annual recreational event for all students in a Japanese elementary school, with students from different grades competing against each other. In other words, the game is actively played within the school. The elementary school where the game takes place has approximately 600 students in total, with around 100 students in each grade from first to sixth. Each participant's time was recorded twice, and the time closest to 60 seconds was included in the data, which was then organised by grade. As per the rules, the average time for each grade was calculated, and the grade with the average time closest to 60 seconds was determined as the winner.

For this study, high school students were given the task of considering the following socially open-ended problem: proposing new rules for the 60-second challenge, based on the raw data collected from the actual recreational event. The central question posed to high school students, which will be discussed in detail later, is whether the rule of using the average time to determine the winner should be maintained for future events. The students were prompted to evaluate the appropriateness of utilising the average as an indicator for determining the winner through a data analysis activity. We believed that engaging in such mathematical activities would contribute to the development of the students' critical thinking skills.

Detailed rules of the 60-second challenge are presented in **Figure 2**.

<60-sec challenge>
All students, first to sixth grades
(1) The student starts the stopwatch and stops it when they think 60 sec have passed, without looking at the display (closing their eyes).
(2) Decimals are rounded to the nearest second, and the data are taken as integer values.
(3) Each student has two tries, of which the closest to 60 sec is included in the data.
(4) If the student fails to record a time on both tries (e.g., if they forget to push the button), the time is recorded as 0 sec.
[Rules for determining the winner]
This is a competition between grades. The grade with the average time closest to 60 sec wins.* Grades that have not learned the concept of average are told that that the grade closest to 60 sec is the winner.

Figure 2. Rules of the 60-sec challenge (Source: Authors' own elaboration)

Average of data for each grade
*Figures in parentheses show the difference from 60 sec.
Roman numerals indicate rank
First grade: 59.59 (-0.41) (ii)
Second grade: 56.71 (-3.29) (vi)
Third grade: 60.42 (+0.42) (iii)
Fourth grade: 59.22 (-0.78) (v)
Fifth grade: 60.68 (+0.68) (iv)
Sixth grade: 59.96 (-0.04) (i)

Figure 3. February 2020 60-sec challenge average results by grade (Source: Authors' own elaboration)

Average of data for each grade
*Figures in parentheses show the difference from 60 sec.
Roman numerals indicate rank
First grade: 55.53 (-4.47) (v1)
Second grade: 58.35 (-1.65) (v)
Third grade: 58.74 (-1.26) (iv)
Fourth grade: 60.29 (+0.29) (iii)
Fifth grade: 60.102 (+0.102) (ii)
Sixth grade: 60.1 (+0.1) (i)

Figure 4. February 2021 60-sec challenge average results by grade (Source: Authors' own elaboration)

Task
The 60-sec inter-grade challenge between grades is held at K Elementary School. Are the existing rules for this challenge appropriate as they currently stand? Review and propose new rules for the 60-sec inter-grade challenge. In doing so, provide an explanation for both teachers and elementary school students (e.g., the basis for why the rule is considered to be good).

Figure 5. Central task for proposing new rules for the 60-sec challenge (Source: Authors' own elaboration)

The 60-second challenge occurred twice in an elementary school, with data gathered per class. Grades 1 and 2 were assisted by teachers. The competition ran in February 2020 and 2021, resulting in two datasets being collected (one per year). Each year, an average time per grade was calculated. In 2020, 6th graders won, followed by 1st graders (**Figure 3**). In 2021, 6th graders won again, followed by 5th graders (**Figure 4**).

The high school students were subsequently given the following task to work on, utilising the provided results (**Figure 5**). **Figure 5** presents the key question that the students were required to contemplate to showcase their critical thinking skills in this experimental class.

On the worksheet, the high school students were instructed to provide two explanations regarding the new rules for the 60-second inter-grade challenge conducted at K Elementary School. One explanation was intended for the teachers, while the other was meant for the elementary school students. It is important to note that the 60-second challenge is not a hypothetical task, but an actual event held at an elementary school. Therefore, the objective was for the high school students to contemplate what kind of rules would be appropriate if the challenge were to be held the following year at K Elementary School. Moreover, since our aim

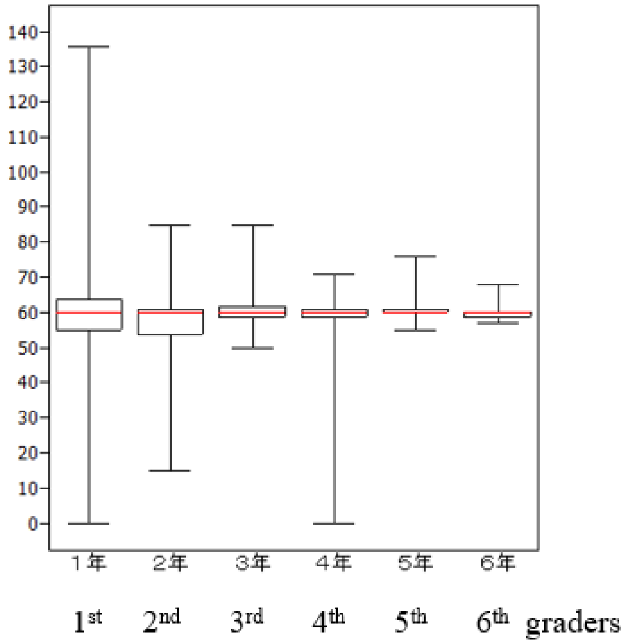


Figure 6. Box-and-whisker diagram for February 2020 60-sec challenge results (Source: Authors' own elaboration, using Simplebox)

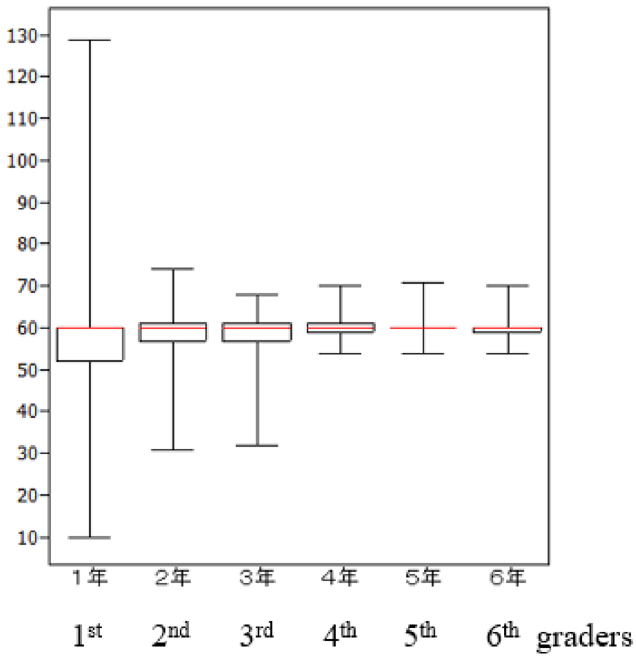


Figure 7. Box-and-whisker diagram for February 2021 60-sec challenge results (Source: Authors' own elaboration, using Simplebox)

in this study is to foster critical thinking in a strong sense, we also wanted to instil a sense of otherness in the students. As a result, they were asked to propose new rules to both the teachers and students at K Elementary School[1].

During the lesson, raw data in the form of 60-second challenge records for each grade were distributed to the students. They utilised statistical analysis tools such as Simplebox (for box-and-whisker plots) and Simplehist (for frequency distribution tables and histograms) on their tablets to analyse the data statistically. For instance, **Figure 6** and **Figure 7** illustrate the representation of the raw data through box-and-whisker plots.

The planned lesson spanned a duration of three hours. In the first hour, students individually considered new rules and then shared their ideas within their respective groups. In the second hour, groups engaged in an information exchange session to share their progress. Finally, in the third hour, each group was given the opportunity to present their newly proposed rules.

[1] After the experimental lesson, the teacher evaluator from K Elementary School proceeded to share the recommendations with the elementary school students.

Table 2. Proposal for weighting by grade

Time range	Points (per person)
Under 50 sec.	All grades: 0 points
50-under 55 sec.	Grades 4–6: 10 points Grades 1–3: 15 points
55-under 60 sec.	Grades 4–6: 15 points Grades 1–3: 20 points
60 sec.	Grades 4–6: 20 points Grades 1–3: 30 points
60-under 65 sec.	Grades 4–6: 15 points Grades 1–3: 20 points
65-under 70 sec.	Grades 4–6: 10 points Grades 1–3: 15 points
70 sec. or more	All grades: 0 points

Table 3. Evaluation by rubric

S (Satisfactory)	A (Excellent)	B (Adequate)	C (Insufficient)
The response satisfies criteria A and includes excellent points, such as a clear rationale presented from multiple perspectives and efforts made to ensure comprehension by different audiences, including teachers and elementary school students.	The rationale behind why the rule is considered good is clearly explained, encompassing the basis for acceptability to a wide range of individuals.	There is some explanation of the rationale behind why the rule is considered good.	No statement or the statement contains major errors.

Handling of socially open-ended problems

As outlined by Shimada (2017), the handling of socially open-ended problems specifically involves utilising mathematical models to address the problems. In our lesson, the task given to the students serves as a mathematical model. Therefore, the question arises: what kind of mathematical model will the students propose?

Based on the data from February 2020, where the first-grade students secured second place, it becomes evident that the rule determining the winner, which is based on the average value, can be reevaluated. As depicted in **Figure 6** and **Figure 7**, some students may not be content with the fact that the first grade, with a wider range of results compared to the other grades, attained second place. Conversely, if students prioritise the value of kindness, they may perceive no issue with this outcome. In such a case, the use of the average value to determine the winner would be upheld.

Additionally, students might propose counting the number of students who achieve exactly 60 seconds. However, since this proposal does not account for the number of students in each grade, it would be more suitable to divide the count of students at exactly 60 seconds by the total number of students in the grade. The latter proposal demonstrates a higher level of mathematical sophistication compared to the former. Furthermore, if the value of kindness towards the first graders is further emphasised, students may suggest implementing weighting. For instance, as illustrated in **Table 2**, a proposal could involve introducing a scoring system for each time range and assigning greater weight to scores for first-to third graders.

Other possibilities include the notion of setting the time ranges with greater precision or modifying them to determine the winner based on the number of individuals within a range of 60 seconds ± 2 seconds. Another possibility is to calculate the average while excluding outliers. As a result, students are encouraged to construct mathematical models that align with their values. Therefore, we anticipate that the responses will differ based on the students' values. This type of lesson exemplifies an environment where students themselves develop creative mathematical activities, allowing them to showcase their own critical thinking abilities through these activities.

Evaluation by rubric

Based on the discussion of evaluation in Hattori et al. (2023), we have developed a rubric (evaluation criteria) which is presented in **Table 3**. When creating the rubric, we placed emphasis on the rationale behind setting the rules. This is because we believe that the rationale determines critical thinking, which is rooted in social values.

The authors view judgments and the presentation of the rationale for them in the following manner. As mentioned earlier, values are principles that guide human behaviour (cf. Sagiv et al., 2017), and they serve as criteria for judgment and a framework for thinking (cf. Morioka et al., 1993). However, in most cases, the criteria used for judgment are not only unobservable but also unrecognised by the individual making the judgment. Therefore, in this study, we focused on how the basis for judgment is presented. This relates to the infinite retroactive structure of the grounds for judgment. In other words, the basis of a judgment itself becomes a judgment, and the basis of that judgment can be further questioned. In principle, the grounds for judgment can be traced back indefinitely. However, in real communication, we do not go back infinitely. Instead, we stop at a point that can be stated indefinitely in principle: 'in this situation, it would be sufficient to state this much'.

Judgment implies the criteria/framework used, as shown by the rationale provided. The rubric in **Table 3** interprets how rationales are presented to understand the values guiding learners' judgments and their alignment with a universal rationale. For instance, if the 'Explanation for Teachers' acknowledges differences in children's developmental stages, it can be inferred that decisions are made with a focus on fairness. Furthermore, if specific differences in children's developmental stages from first to sixth grades are mentioned, such as variances in the perception of time quantity, then it suggests a broader rationale accepted more widely. This rubric helps assess how social values are manifested.

These rubric prompts student to offer clear explanations that consider individuals with different perspectives. This approach was adopted to foster the development of critical thinking in a strong sense. Critical thinkers in the strong sense are expected to consider the interests of diverse individuals and groups beyond their own.

MATERIALS AND METHODS

The class was held at a high school affiliated with a national university, with approximately 200 students per grade (five classes of 40 students each). Of these students, approximately 60% entered the school from the junior high school affiliated with the university. Classes are not organised based on proficiency levels, and there is a mix of students who are extremely good at math and those who feel that math is not their strong suit. However, compared with a typical Japanese school, this school has a more academically advanced student body, and almost all students can perform basic calculations that can be found in textbooks. This experimental lesson was conducted in a second-year high school class including 40 students, comprising 23 male and 17 female students. The lesson was implemented over three consecutive days in late June 2021, with each session lasting 45 minutes. The instructor oversees 'Math II' for this class and teaches four hours every week. He is also the homeroom teacher of the class and communicates well with the students. This is the first time that a class designed to foster critical thinking skills has been implemented.

Data were collected through video recordings, group reports on worksheets, and questionnaires. A camera was placed at the back of the classroom to capture an overall view of the class. During group activities, an additional handheld camera was used to record conversations by positioning it close to each group while filming.

During the experimental class, an information exchange session was conducted in the second period, followed by a peer evaluation of other groups' presentations in the third period. These activities were designed to foster critical thinking in a strong sense, as outlined in the theoretical framework. The aim was to provide students with ample opportunities for interaction and to emphasise the importance of sharing ideas with others. The decision to schedule lessons over three consecutive days was made to emphasise interactions with others. Critical thinking in its strong sense involves receiving the opinions of others and reconsidering one's own perspectives. Particularly, during the information-sharing session on the second day, ample time was allocated for students to listen to others' opinions. This approach was intended to encourage students to constructively reflect on their views and make fair proposals based on these reflections.

After the class, the worksheets were collected, scanned, and analysed. Three rubric evaluators were involved: the high school teacher who delivered the lesson (the second author), a university teacher (the first author), and an elementary school teacher who conducted the new 60-sec challenge game. During this lesson, the students were expected to present diverse opinions from different perspectives. Consequently, the analysis of the content written on the worksheets inevitably depends, to some extent, on evaluators' value judgments. To prevent any evaluator from being influenced by the opinions of others, three evaluators independently conducted the analyses and rubric-based evaluations. Specifically, each evaluator worked separately, not at the same location, to qualitatively assess the worksheets individually, based on **Table 3**. Scale S of the rubric was specifically focused on the criterion of 'Steps taken to enable comprehension by different audiences, namely, teachers and elementary school students'. Since the students' reports were proposals for new rules for the 60-sec challenge, it was deemed appropriate to have an elementary school teacher from the corresponding school evaluate the reports and provide sincere feedback to the students. The expected outcome of the students' critical thinking was the proposal of an alternative rule to replace the mean. The worksheet descriptions were examined to identify the values emphasised by the students, and the demonstrated critical thinking was evaluated. Subsequently, a questionnaire was administered during the final 10 minutes of the third period, followed by a text mining analysis. This questionnaire was designed to measure the extent to which students demonstrated critical thinking in its strong sense. As previously mentioned, this practice aims to encourage students to critically evaluate the rules proposed by others and revise their own rules for improvement. In other words, the questionnaire sought to gauge how consciously students considered others' opinions. By employing text mining techniques, we analysed the keywords provided in the responses to identify the relationship between the improvement of students' own ideas and the influence of others' perspectives.

In the next section, we will describe and discuss the classroom practice, focusing on three perspectives: the mathematics used by students in the class, the functioning of the rubric, and the demonstration of critical thinking by the students.

LESSON IMPLEMENTATION AND DISCUSSION

Lesson implementation

First session

First, the progression of the three-hour lesson was explained, and a worksheet was distributed to the students. The teacher and students then collectively established the task. A time limit of 3 minutes was set, during which the students individually contemplated new rules for the 60-sec challenge. Each group, comprising five members, received a tablet PC. The group discussions were to be conducted using the tablets, and the operation of the computer software was explained to the entire class prior to the activity. The students were informed that they could utilise SimpleHist and Simplebox for analysis and were permitted to modify the original data (e.g., considering outliers) or download additional software if needed. The tablets provided internet access, and students were allowed to perform searches, among other activities. Furthermore, the use of smartphones was also allowed. During the group discussions, students were required to utilise the tablets to review data and experiment with the 60-sec challenge using their smartphones.

In the first session, attention was focused on rationality, as defined in this study's critical thinking framework. This session involved activities in which students critically examined existing rules and explored the possibility of creating better alternatives.

Second session

After explaining the session's progression at the beginning, students were instructed to assign roles for the information exchange activity (7 minutes each for two rounds)-deciding who would present in the first half and second half and assigning roles for listening. During the information exchange, the listeners took on the roles of either a teacher or an elementary school student and provided feedback to the presenter. For instance, a student playing the role of an elementary school student asked, 'What is an absolute value?' to highlight the importance of terminology.

The following are examples of actual student statements made during the information exchange activity:

> Regarding the rule for determining the winning grade, if we simply calculate the average, a situation where half of the participants score 0 seconds, and the other half score 120 seconds would result in an average of 60 seconds. Even though these scores are not close to 60 seconds, the average gives them a high evaluation. To address this issue, I think it would be better to take the absolute difference from 60 seconds.
>
> If we calculate the average, for example, a case with two participants scoring 90 seconds and 30 seconds would produce the same result as two participants both scoring 60 seconds. To improve this, I considered assigning points based on how far each score is from 60 seconds and designing the game around that idea.

These examples demonstrate that both groups of students identified the shortcomings of using averages and proposed the mathematical concept of variance as a solution. Through this information exchange within each group, students sought to refine their ideas and explore more appropriate mathematical approaches.

Following the information exchange, each group was tasked with completing their reports. Four out of the eight groups submitted their reports during the lesson, while the remaining groups were instructed to submit their reports on the same day.

In the reports of each group, the following ideas were observed:

1. Rounding off the decimal part of the time instead of truncating it (or including the decimal part in individual records),
2. Implementing handicaps based on grade, and
3. Devising methods to prevent outliers from affecting the data.

The first point was raised by one of the groups during their presentation in session 3, emphasising that a time of 59.9 seconds should be considered worse than 60.4 seconds. Regarding the second point, Group B proposed a rule where the larger the deviation from 60 seconds, the lower the score. They set specific handicaps for each grade, such as deducting 1 point for every 10-second deviation in the range of 40-50 seconds and 50-60 seconds for 1st graders and deducting 1 point for every 10-second deviation in the range of 58-59 seconds and 59-60 seconds for 6th graders, and so on. When determining the number of seconds that should result in a point deduction for each grade, the students aimed to maintain a similar shape in the histogram as the time range changed. Though their approach was intuitive, their discussion had a mathematical foundation.

Regarding the third point, Group C proposed transforming the competition into a relay format, where groups of three people took turns counting 20 seconds each. The students noted that this method would reduce the number of outliers as one player could correct another's mistakes. Additionally, it addressed the issue of boredom that arose from staying quiet for 60 seconds, emphasising the value of enjoying the game.

The second session focused on activities involving information exchange among the students. Criticality, as a component of critical thinking, is expected to be demonstrated through a critical examination of rules created by others. Furthermore, the session was designed to encourage students to reflect on their own rules based on others' opinions and to strive to employ better mathematics skills. Therefore, a demonstration of 'otherness' was anticipated in this session.

Third session

In the third session, each group had three minutes to deliver their presentation, and a projector was available for their use. During the presentations, students were encouraged to conduct a brief peer evaluation of the other groups. Following the presentations, a 10-minute period was allocated for participants to complete a questionnaire. The peer evaluations were scored on a scale of 0-10, with each student awarding their own group a score of 7. However, students were not required to provide a rationale for their scores, as the purpose of this exercise was to foster a sense of competition among the students. The scores awarded to each group are shown in **Table 4**.

Feedback

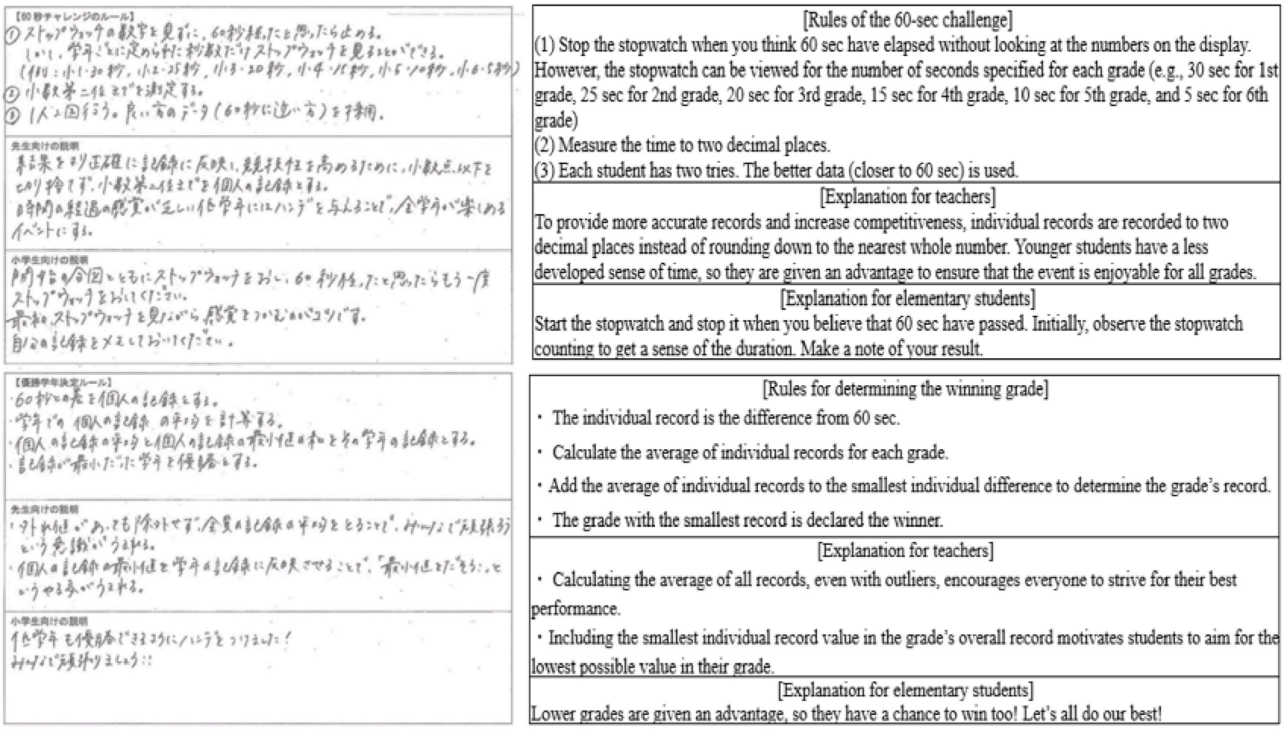
Subsequently, the students were provided with a feedback sheet that included comments from the three evaluators mentioned earlier. These evaluations were conducted based on the rubric presented in **Table 3**. The evaluation results from the three teachers (T1, T2, and T3) are presented in **Table 5**.

Table 4. Scores for each group (evaluated by students)

	A	B	C	D	E	F	G	H
	7	7	10	9	9	8	8	6
	6	7	4	7	6	6	5	5
	6	7	7	6	6	7	6	6
	6	8	7	7	8	6	7	5
	6	9	8	7	7	8	7	7
	4	5	6	5	5	7	4	4
	8	9	8	10	9	8	7	8
	4	7	8	9	9	8	7	7
Total	47	59	58	60	59	58	51	48

Table 5. Evaluation results by the three teachers (T1, T2, and T3)

Group	T1	T2	T3
Group A	B	B	B
Group B	A	B	B
Group C	S	S	S
Group D	A	A	A
Group E	S	A	S
Group F	A	A	S
Group G	B	B	A
Group H	A	A	S



【60秒チャレンジのルール】
①ストップウォッチの数字を見ずに、60秒経ったと思ったら止める。
しかし、学年ごとに定められた秒数だけストップウォッチを見ることができる。
（例：小1・30秒，小2・25秒，小3・20秒，小4・15秒，小5・10秒，小6・5秒）
②小数第二位までを測定する。
③1人2回行う。良い方のデータ（60秒に近い方）を採用。

先生向けの説明
結果をより正確に記録に反映し、競技性を高めるために、小数点以下を切り捨てず、小数第二位までを個人の記録とする。
時間の経過の感覚が乏しい低学年にはハンデを与えることで、全学年が楽しめるイベントにする。

小学生向けの説明
開始の合図とともにストップウォッチをおし、60秒経ったと思ったらもう一度ストップウォッチをおしてください。
最初はストップウォッチを見ながら感覚をつかむのがコツです。
自分の記録をメモしておいてください。

【優勝学年決定ルール】
・60秒との差を個人の記録とする。
・学年での個人の記録の平均を計算する。
・個人の記録の平均と個人の記録の最小値の和をその学年の記録とする。
・記録が最小だった学年を優勝とする。

先生向けの説明
・外れ値があっても除外せず、全員の記録の平均をとることで、みんなで頑張ろうという意識がうまれる。
・個人の記録の最小値を学年の記録に反映させることで、「最小値をだそう」というやる気がうまれる。

小学生向けの説明
低学年も優勝できるようにハンデをつけました！
みんなで頑張りましょう！！

[Rules of the 60-sec challenge]
(1) Stop the stopwatch when you think 60 sec have elapsed without looking at the numbers on the display. However, the stopwatch can be viewed for the number of seconds specified for each grade (e.g., 30 sec for 1st grade, 25 sec for 2nd grade, 20 sec for 3rd grade, 15 sec for 4th grade, 10 sec for 5th grade, and 5 sec for 6th grade)
(2) Measure the time to two decimal places.
(3) Each student has two tries. The better data (closer to 60 sec) is used.

[Explanation for teachers]
To provide more accurate records and increase competitiveness, individual records are recorded to two decimal places instead of rounding down to the nearest whole number. Younger students have a less developed sense of time, so they are given an advantage to ensure that the event is enjoyable for all grades.

[Explanation for elementary students]
Start the stopwatch and stop it when you believe that 60 sec have passed. Initially, observe the stopwatch counting to get a sense of the duration. Make a note of your result.

[Rules for determining the winning grade]
- The individual record is the difference from 60 sec.
- Calculate the average of individual records for each grade.
- Add the average of individual records to the smallest individual difference to determine the grade's record.
- The grade with the smallest record is declared the winner.

[Explanation for teachers]
- Calculating the average of all records, even with outliers, encourages everyone to strive for their best performance.
- Including the smallest individual record value in the grade's overall record motivates students to aim for the lowest possible value in their grade.

[Explanation for elementary students]
Lower grades are given an advantage, so they have a chance to win too! Let's all do our best!

Figure 8. Report for Group D (Source: Field study)

The evaluations were conducted independently by the teachers based on the reports submitted by the students at the end of the third session. A comparison of **Table 4** and **Table 5** reveals that student and teacher evaluations do not necessarily align. We present the students' evaluation of Group D (which received the highest score, as shown in **Figure 8**) and the teachers' evaluation of the same group (as shown in **Table 6**). These comments were shared with the students. **Tables 6-11** display additional evaluation comments provided by the teachers.

In **Figure 8**, Group D includes the statement, 'Lower grades are given an advantage, so they have a chance to win too!' This can be interpreted as an expression of fairness, reflecting the value of empathy towards younger grades.

Table 6. Evaluation comments given by the teachers regarding Group D's report

Evaluation	Comment
A	The group's consideration of the handicap shows their intention to ensure that even younger students can enjoy the game. The explanation provided to elementary school students about observing the stopwatch count to get a sense of time reflects their thoughtful approach. However, I felt that the explanation regarding the appropriateness of the handicap could have been clearer (e.g., Is the handicap perfectly balanced?). The group's presentation had a strong presence that seemed to resonate with the elementary school students.

Table 6 (Continued). Evaluation comments given by the teachers regarding Group D's report

Evaluation	Comment
A	Regarding the rules of the 60-sec challenge, the group came up with a method to provide an advantage to the younger students. The warmth and helpfulness of the group members were evident in their explanation to the elementary school students. The phrase "the sum of the average of individual records and the smallest individual difference" could have different interpretations, and it would be helpful to provide specific examples for better understanding.
A	I appreciate how the rules were designed to make the event enjoyable for all grades by giving younger students more time to observe the stopwatch. The idea of recording the sum of the minimum and the mean as the grade's record is excellent, as it motivates students to aim for the minimum value and maintains excitement until the end. It would be even better if the group could effectively convey the fun of this aspect in a way that is easily understandable for elementary school students.

Table 7. Evaluation comments given by the teachers regarding Group B's report

Evaluation	Comment
A	Group B's explanation of the appropriateness of the handicap based on the shape of histograms when changing grades was impressive. It's unfortunate that this aspect wasn't verbalised during the presentation and reflected in the report. The inclusion of prizes in the rules aligns with reward-oriented motivation models, and it's noteworthy that motivation was sought beyond the game itself.
B	The rules of the 60-second challenge were designed to minimise mistakes, such as forgetting to press the button, by allowing only 3 seconds for the first attempt. The inclusion of a reward, like the *momiji manju* (maple-shape manju cake), for determining the winner added excitement to the competition. However, it would have been helpful to have a more explicit explanation of why these rules were devised in their specific ways.
B	The idea of assigning different point ranges for each grade level in the 60-second challenge is quite interesting. To help elementary school students better understand this concept, using the term "range of ranks" in the explanation could be beneficial. Additionally, providing clearer rationales and recommendations for each rule would enhance understanding.

Table 8. Evaluation comments given by the teachers regarding Group E's report

Evaluation	Comment
S	The explanation provided to the teachers regarding recording up to the second decimal place was rooted in the shortcomings of previous rules. On the other hand, the explanation given to the elementary school students was more intuitive and well-crafted. The contestants demonstrated a strong awareness of making the game engaging for elementary school students when determining the winning grade. They devised various prizes and carefully considered which grade would have the best chance of winning. I found this thoughtful consideration based on real-world situations to be commendable.
A	The rules for the 60-second challenge effectively utilise the stopwatch data as it is. The intention behind these rules was well explained and presented convincingly. Additionally, the rules for determining the winner, such as incorporating the use of absolute values, were cleverly devised. However, it is worth noting that the explanation provided for elementary school students may have been slightly challenging to comprehend.
S	The idea of utilising a record to two decimal places is highly persuasive, as it prevents potential score reversals that could occur with integer values (e.g., 60.99 and 59.99). The rule of awarding bonus points to the child who surpasses the 60-second record or the 59.99 record among all the students adds an intriguing element to the challenge. Additionally, the concept of assigning names to prizes, such as the "prize for everyone working together," is a fantastic idea that enhances both the understanding and enjoyment of the activity for elementary school students.

Table 9. Evaluation comments given by the teachers regarding Group F's report

Evaluation	Comment
A	It would be beneficial to provide more elaboration on the appropriateness of the handicap, such as discussing whether the designated counting time is optimal or if it would be better to divide the scoring into different grade ranges (1st–2nd / 3rd–6th). A comparison of histograms and the number of data satisfying specific conditions could serve as compelling evidence. The names "60-second Meister" and "title" are engaging for elementary school students, and it is evident that the target audience was well-considered in the design of the rules.
A	The proposed new rules for the 60-second challenge are highly convincing, as they address the difficulties faced by younger students in the game. The rules for determining the winner are also clever, particularly the inclusion of percentages and the introduction of a separate box. Providing more specific and concrete figures, such as **Figure 1** and **Figure 2**, would have made the explanations easier to understand for both teachers and elementary school students.
S	I find the concept of providing a handicap by counting to the halfway point based on the grade to be quite appealing. The names "60-second Meister" and "semi-60-second Meister" are also creative and engaging, making it easier for children to grasp the concept and adding excitement to the competition. I am curious to see how these ideas would unfold in practice.

Table 10. Evaluation comments given by the teachers regarding Group G's report

Evaluation	Comment
B	While I appreciate the intention of creating a fun atmosphere, it would be beneficial to discuss whether the handicap is suitable for achieving this goal. The idea of eliminating homework aligns with reward-oriented motivation models, and it is commendable to seek motivation beyond the game itself. However, the explanation provided for elementary school students seems too simplistic. It would have been helpful to invest more effort in explaining the rules of the "60-second challenge" in greater detail.
B	The rules for the 60-second challenge were designed to provide an opportunity for students in lower grades to compete on a more equal footing using a handicap. The rule for determining the winner, using mathematical calculations based on the average of absolute values, demonstrates careful thought and consideration. It would have been advantageous to provide clearer explanations of the rules for the 60-second challenge and the criteria for determining the winning grade to elementary school students.

Table 10 (Continued). Evaluation comments given by the teachers regarding Group G's report

Evaluation	Comment
A	The concept of adjusting the number of measurements in the handicap idea is intriguing. However, it might have been more exciting if the focus was on devising game-related rules to enhance the excitement, rather than offering a reward unrelated to the game, such as the elimination of homework. I appreciate the idea of comparing the average difference (absolute value) from 60 rather than the average times, as it aligns with the game's objective of having the participant closest to 60 seconds emerge as the winner.

Table 11. Evaluation comments given by the teachers regarding Group H's report

Evaluation	Comment
A	The explanation provided for rounding off instead of rounding down was carefully articulated, highlighting the advantages of this approach. The argument that the modification allows those who struggle with the 60-second challenge to aim for lower times while providing an opportunity for those skilled at the challenge to aim for higher times was effectively conveyed. The proposed rule to determine the grade winner based on the sum of differences from 60 seconds, however, may disadvantage grades with a larger number of participants. Given more time for rule consideration, a more balanced and equitable rule could have been formulated.
A	The issues with the current 60-second challenge rule, particularly the problems arising from rounding down, were clearly addressed and improved upon. In terms of determining the winner, the use of the absolute value of the difference was introduced. However, it is important to consider the discrepancy in the number of students across different grades when calculating the sum. If the rationale for not accounting for this difference is explained further, it would strengthen the proposal.
A	The idea of rounding off to one decimal place is highly persuasive, as it resolves the issue of reversing scores like 60.99 and 59.99. The explanation for elementary school students is well-crafted to ensure easy comprehension. The focus on comparing the average difference from 60 rather than the average times is straightforward, supported by clear examples. Introducing a handicap to foster fair competition between lower and upper grade students could add an extra level of excitement to the challenge.

Additionally, Group E provides the following statement:

> Rather than using integers, data should be recorded up to the second decimal place. Reason: If 60.99 and 59.99 are rounded to integers, 60.99, which is farther from 60 seconds, would receive a higher score, which is unfair.

This proposal aims to address unfair situations arising from the traditional rule of using integers for evaluation. This can be considered an explicit manifestation of the value that the game should be made fairer.

The third session involved peer evaluation activities among the students. This session was aimed to demonstrate criticality and otherness as components of critical thinking. Students were required to critically evaluate the rules created by their peers.

Lesson Discussion

In this section, we will analyse the lesson from three perspectives: the students' perspective in proposing new indicators, rubric functioning, and the critical thinking demonstrated by the students.

Students' perspective in proposing new indicators

First, we provide an overview of the improvement proposals from the eight groups. The proposed improvements can be primarily classified into three categories:

1. New implementation methods that do not involve handicaps, such as allowing retries or forming groups of three,
2. New implementation methods that consider differences in ability between grades and include handicaps, such as limiting the number of attempts or adjusting the measurement time, and
3. Other methods. Furthermore, the proposed improvements for determining the winning grade can be classified into seven categories:
    a) Considering the issue of rounding down to the first decimal place.
    b) Taking outliers into account.
    c) Addressing the problem of mean evaluation by considering absolute deviation.
    d) Setting and weighting the intervals.
    e) Differentiating the weighting of intervals for different grades.
    f) Adjusting the interval width for each grade.
    g) Introducing special points, like an award for a perfect guess of 60 seconds.

The classification was based solely on the explicit statements found in the students' written worksheets, without incorporating any assumptions made by the authors. In this section, we will analyse the report provided by Group D (**Figure 8**).

Based on Group D's improvement plan, we can classify it as Category II due to the following description found in the worksheet:

> 'The younger students have less sense of passage of time, so they are given an advantage in a way that the event can be enjoyed by all grades'.

Table 12. Student perspectives and suggested improvements to indicators

Group	Classification
Group A	I / a, b, c, d
Group B	III / c, d, f
Group C	I /a, b, d
Group D	II / a, c
Group E	III / a, c, g
Group F	II / a, d, f, g
Group G	II / a, c
Group H	III / a, c

This suggests an implementation method that sets up a handicap to favour the lower grades by introducing a difference in the measurement time for each grade. Regarding the indicator used to determine the winning team, Category a is identified from the description: 'To reflect the results more accurately in the records and to enhance competitiveness, individual records are recorded to two decimal places rather than rounding down to the nearest whole number'. Additionally, Category c is identified from the description: 'The difference from 60 seconds will be considered as the individual record'. Furthermore, the statement 'Lower grades are given an advantage, so they have a chance of winning too!' can be interpreted as an expression of the value of 'kindness toward the first graders'. It signifies that the proposal aims to provide an opportunity for lower-grade students to compete on an equal footing. In terms of values, there are other potential values that emphasise the degree of scattering. The proposed improvements from each presentation are summarised in **Table 12**.

Among several groups, the issue of adopting the average value (category c) resulted in considering the difference from 60 seconds. This approach focuses on the dispersion of data, highlighting the importance of considering variance, standard deviation, and other measures of variability among students, rather than solely relying on the mean as a central indicator. Additionally, various methods for weighting the intervals based on differences in ability across grades were observed.

Particularly noteworthy was a group that approached the problem from the perspective of adjusting the interval widths (category f). They created histograms at different interval widths and aimed to transform the distributions to be as similar as possible. This approach demonstrated their unique way of addressing the problem.

The data for each grade exhibited typical characteristics, with minimal variation in means but significant differences in scatter. Variance and standard deviation are commonly used to measure scatter, and in traditional teaching settings, teachers often present data with similar characteristics and set up specific scenarios to prompt students to consider the degree of dispersion. However, in this case, the students recognised the importance of dispersion based on their own values (although not all students may have shared this perspective). While adjusting the interval widths, the students might have applied the concept of standardisation, seeking to achieve similar distributions to maintain a competitive event. This teaching assignment stands out as it involves setting competition rules based on students' values and witnessing the spontaneous use of standardisation and similar approaches.

Furthermore, a group proposing an improved implementation method (category II) that accounted for differences in ability displayed an awareness that increasing the number of attempts led to reduced scatter in the average values. The students likely had an intuitive understanding of the characteristics of the sample mean distribution. Considering these circumstances, the students put forth new indicators based on their own values and actively utilised important statistical concepts learned in secondary education.

In this way, students utilised mathematics to a greater extent than anticipated. This was driven by the value of empathy, as they aimed to make the activity enjoyable, even for younger students. Notably, their focus on dispersion aligns with contemporary mathematical literacy, as emphasised by the Organisation for Economic Co-operation and Development (2018). The context provided by these teaching materials facilitates the development of mathematical literacy.

Rubric functioning

Eight group reports were evaluated by three teachers using the rubric from **Table 3**, showing no significant discrepancies (**Table 5**). For example, Group C was rated 'S' by all for transforming the challenge into a team game, setting winning criteria, removing outliers, and creating an enjoyable experience. Group D, highest-rated by students, was awarded 'A' for adjusting the rules by providing more time for lower grades to view the stopwatch and fostering motivation. Despite diverse evaluation rationales from teachers, Group D appeared appropriate from one perspective and challenging to understand mathematically from another. Similar evaluation approaches were employed for the other groups using the same rubric. These findings demonstrate that the rubric-based assessment presented in **Table 3** serves as a tool to assess the social values demonstrated by the students. However, it may be exceptionally difficult to evaluate cases where one perspective has a highly compelling rationale while another perspective is incorrect based solely on this rubric. Therefore, it is crucial to further refine the evaluation criteria for the rubric.

Moreover, evaluators' judgments may differ due to their backgrounds. Analysing the ratings given by the three teachers, it can be observed that some differences may be attributed to their professional backgrounds. For instance, T1 was a classroom teacher, T2 was a university teacher, and T3 was an elementary school teacher. While the evaluations of the three teachers were never more than two levels apart, T3 generally provided higher ratings compared to T2 and T1 (e.g., **Table 9**, **Table 10** and **Table 11**). The teacher evaluation comments in **Table 10** are particularly interesting. Both T1 and T2, who assigned a B rating, requested more detailed explanations suitable for elementary school students, as seen in their comments such as 'it would have been

1. In this lesson, you came up with rules for determining the winner of the 60-sec challenge. At the beginning of the second period, after hearing or seeing your friends' rules for determining the winner and their rationale, did you decide to change your own rules for determining the winner? Please tick one of the following. Also, please give the reason (if you decided to change your rule, give the reason or reasons for the change; if not, give the reason or reasons).
2. What skills do you think you acquired in this lesson? Please state what skills you gained and why you thought so.
3. Please feel free to describe your impressions of this lesson.

Figure 9. Post-lesson questionnaire (Source: Authors' own elaboration)

Table 13. Extracted feature words: Top ranking words

Decided to change the rule		Decided not to change the rule	
feel	.307	rule	.250
opinion	.165	oneself	.230
good	.136	group (Han)	.192
interesting	.107	think	.191
grade	.104	enjoy	.106
Elementary school student	.095	person	.100
presentation	.086	mathematics	.094
listen	.080	group	.078
see	.078	ability	.068
game	.078	other	.067

English translation of the results of text mining for Japanese descriptions

advantageous to provide clearer explanations of the rules for the 60-second challenge and the criteria for determining the winning grade to elementary school students' and 'However, the explanation provided for elementary school students seems too simplistic'. Conversely, T3, who assigned an A rating, highly evaluated the concept of the new rule itself, stating, 'The concept of adjusting the number of measurements in the handicap idea is intriguing'. T3 is an elementary school teacher, whereas T1 and T2 are not. T3 considered the explanation adequate for elementary school students. By contrast, T1 and T2, who do not regularly interact with elementary school students, may have set a higher standard for what constitutes 'clarity' in explanations for this age group. These variations in evaluation priorities highlight the differing perspectives of the evaluators. Consequently, in some studies, rubric evaluation criteria are shared among evaluators to establish typical examples of each evaluation in advance. This approach assists in assuming predictable viewpoints and presenting evidence for judgments from these perspectives (Hisadomi & Koyama, 2018).

However, we believe there is an advantage in providing diverse feedback to students through varied evaluations from teachers. In this classroom study, the evaluation results from the three teachers were shared with the students. As shown in **Table 5**, three out of the eight groups received consensus among the teacher evaluations, while five groups did not. When fostering critical thinking skills, it is crucial to emphasise multidirectionality rather than directing abilities in a single direction. It is essential to listen to various ideas presented by others. Therefore, the authors believe that it is important for students to embrace diversity in evaluations. Hence, both the evaluations and the corresponding comments were presented to the students.

Critical thinking demonstrated by students

After the lesson, a questionnaire (**Figure 9**) was distributed, and a text mining analysis was performed using the free software KH coder. Typographical errors were corrected during the analysis.

Regarding question 1, 20 students indicated that they altered their recommendations, while 19 students stated that they did not make any changes. The three free-response statements were combined, and characteristic words were extracted from the responses using the Jaccard coefficient, with the sentence as the unit of aggregation. The extracted words are presented in **Table 13**.

The analysis revealed that the words 'elementary school student', 'grade', and 'game' were prominent in the responses regarding why they changed the rule, indicating that the respondents envisioned elementary school students engaging in the 60 - second challenge (*The Jaccard coefficient for 'rule' in 'decided to change the rule' was 0.190 but filtered out). Here are some representative responses:

> I came up with a better idea by taking on the opinions of others in addition to my own thoughts. There were not only opinions but also various ways of thinking. In this case, there were the concepts of entertaining elementary school students and making the competition interesting, which was interesting.

> I aimed to create explanations that even elementary school students could understand by avoiding technical terms and using simple language.

> I pondered how to make the competition more enjoyable for elementary school students, exploring possibilities like modifying the fundamental rules or turning it into a scoring game.

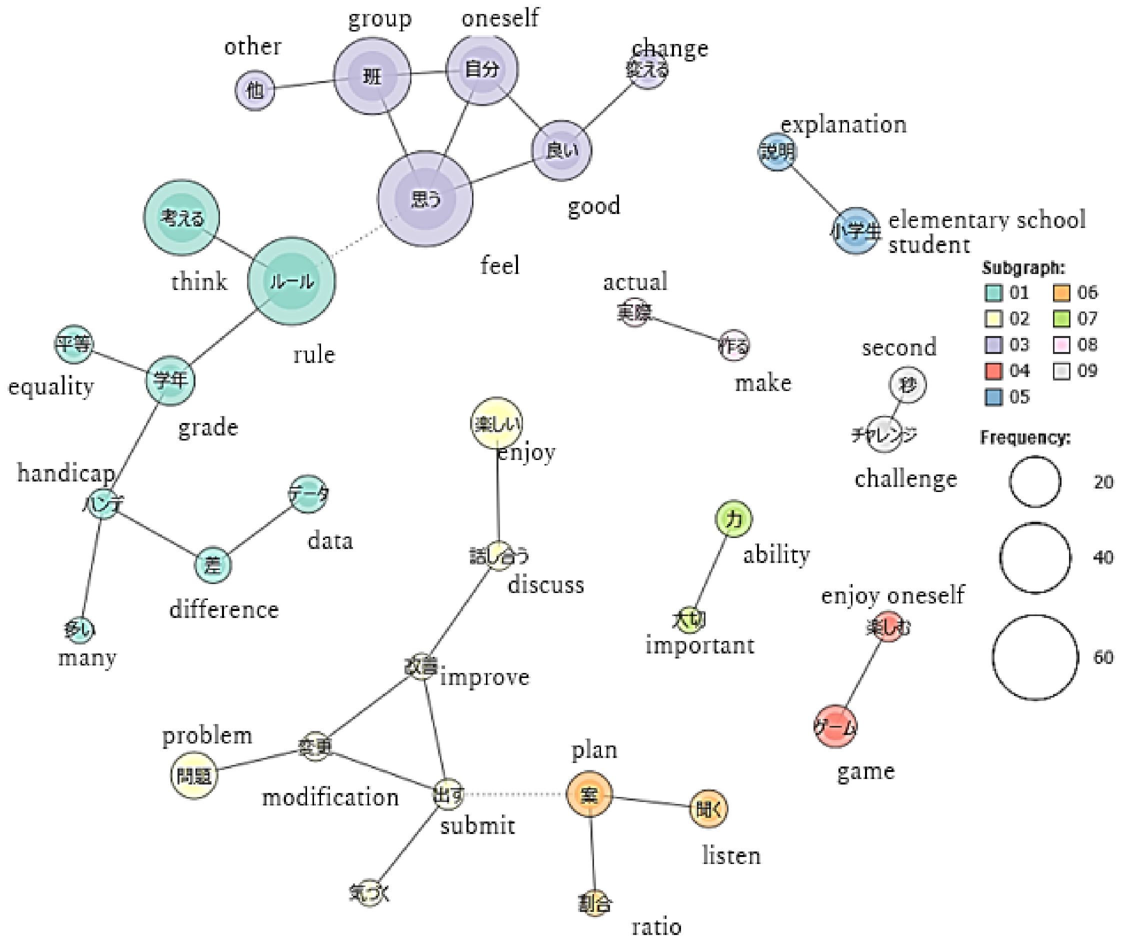


Figure 10. Co-occurrence network (Source: Authors' own elaboration, using KH coder)

These responses demonstrate a strong sense of critical thinking. Throughout the lesson, the high school students exhibited an attitude of seeking to enhance the rules for the enjoyment of elementary school students. This is evident from the questionnaire results, where students actively considered the opinions of others and made efforts to improve their own approaches. The co - occurrence network diagram (**Figure 10**) further illustrates this, showing the co-occurrence of 'discuss' and 'improve'. The diagram is based on sentence-level aggregation, targeting words with a minimum of 5 occurrences, and depicting the top 30 pairs of words with the highest Jaccard coefficient as edges. These skills aligned with the specific critical thinking abilities that we aimed to foster through this lesson, which involves empathising with others' opinions and collaboratively considering ways to improve rules. A student's response, stating, 'It was enjoyable to think, discuss, and gradually improve the rules together', serves as a quintessential example of this process.

During the process of improving the rules, many groups demonstrated critical thinking by devising mathematically sophisticated rules, as previously mentioned. The issues related to adopting averages, truncating data, and addressing outliers cannot be easily identified through a superficial analysis of the data alone. Instead, they require careful mathematical examination from multiple perspectives. Typically, it is the teacher's questioning in the classroom, such as prompting students with questions like 'What do you think about this idea'? that initiates their thinking on these matters. However, in this case, the students were able to uncover these issues through their own initiative. This indicates the active exercise of critical thinking by the students.

The proposed improvement, which stems from the compassionate value of providing an advantage to younger students, utilises mathematics to realise the concept of the 60-second challenge as an enjoyable competition for all students. The fact that we were able to create a classroom environment where various mathematical models (indicators) were developed based on the individual social values, and where students could share their ideas with others, suggests that this lesson to some extent fostered the students' critical thinking abilities. Here, we revisit the demonstration of critical thinking in the narrow and broad senses, as discussed in the theoretical foundation, and summarise the findings. Critical thinking in the narrow sense refers to critical thinking within the mathematical domain, wherein it is essential to utilise mathematical knowledge to explore better rules. In this regard, students demonstrated critical thinking in the narrow sense by evaluating the validity of averages and proposing new ideas, such as focusing on absolute values and variance. In the broad sense, critical thinking refers to the ability to make social judgments about real-world problems using mathematics. This study specifically emphasised respecting others' ideas and revising one's own thoughts. As shown in the text mining analysis, students empathised with the perspective of elementary school students, striving to provide clearer explanations and devise better rules that could be more enjoyable for younger students. These efforts exceeded the authors' expectations and were considered a clear demonstration of critical thinking in the broad sense.

CONCLUSION

The aim of this study, focusing on secondary mathematics education, was to develop teaching materials that promote critical thinking, describe students' problem-solving processes through practical application, and assess the effectiveness of a rubric for capturing students expressed social values. We designed an open-ended problem centred around proposing new rules for the 60-second challenge. Through the structured lesson, we observed how students independently devised new indicators based on their own social values. Additionally, the use of the rubric in a series of assessment activities played a significant role in identifying students' social values. However, we also encountered certain issues, particularly in refining the evaluation criteria of the rubric. The future challenge lies in addressing these issues and developing an evaluation system that effectively captures the students' demonstrated critical thinking, utilising the insights gained from this study as a foundation.

Author contributions: YHA: conceptualization, writing – original draft, writing – review & editing, supervision, project administration, funding acquisition; **YI:** validation, formal analysis, investigation, resources, data curation; **KM, RH & YHI:** methodology, validation, data curation. All authors have agreed with the results and conclusions.

Acknowledgments: The authors would like to express their gratitude to Editage (www.editage.com) for providing English language editing services.

Funding: This work is supported by the 38th (FY 2020) Casio Science Foundation Research Grant, KAKENHI (21K02470, 24K00420).

Ethical statement: The authors stated that, according to applicable institutional regulations, the study does not require approval from an ethics committee since it was conducted as part of the school's regular educational curriculum. The participants were informed about the purpose of the research, their anonymity.Furthermore, all collected data were anonymized, and no personally identifiable information was recorded or disclosed. The study adhered to strict confidentiality protocols, ensuring that participants' identities remained protected throughout the research process. Any student responses or classroom interactions used in the study were processed in a way that prevents individual identification. The data was securely stored and accessible only to the research team for analysis purposes.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Baba, T. (2007). Primary mathematics education in a society and times with multiple values. *Journal of Japan Society of Mathematics Education*, *89*(10), 20-27.

Baba, T. (2009). Analysis of socially open-ended problems in mathematics education from the perspective of values. *Journal of JASME Research in Mathematics Education*, *15*(2), 51-57. https://doi.org/10.24529/jasme.15.2_51

Baba, T. (2016). Nurturing critical thinking competence through mathematics education. *Proceedings of the 4th Spring Research Conference* (pp. 95-96). Japan Society of Mathematical Education.

Baba, T. (2017). Nurturing critical thinking competence through mathematics education: Consideration of critical mathematics education (Skovsmose, 1994) in Japanese context. In *Proceedings of the 5th Spring Research Conference* (pp. 199-200). Japan Society of Mathematical Education.

Baba, T. (2019). Nurturing critical thinking competence through mathematics education: Consideration of commonalities and differences across pre-primary, primary, secondary schools and universities. In *Proceedings of the 7th Spring Research Conference* (pp. 1-2). Japan Society of Mathematical Education.

Baba, T. (2020). Nurturing critical thinking competence through mathematics education: Comprehensive consideration with ideas of exemplars. In *Proceedings of the 8th Spring Research Conference* (pp. 129-130). Japan Society of Mathematical Education.

Baba, T., & Shimada, I. (2019). Socially open-ended problems for enriching student learning with mathematical models and social values. In P. Clarkson, W. T. Seah, & J. Pang (Eds.), *Values and valuing in mathematics education: Scanning and scoping the territory* (pp. 171-183). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-16892-6_12

Benedict, R. (1946). *The chrysanthemum and the sword; patterns of Japanese culture.* Houghton Mifflin.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron, & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9-26). W H Freeman/Times Books/ Henry Holt & Co.

Fonseca, L., & Arezes, S. (2017). A didactic proposal to develop critical thinking in mathematics: The case of Tomás. *Journal of the European Teacher Education Network 12,* 37-48.

Frankenstein, M. (2012). Beyond math content and process: Proposals for underlying aspects of social justice education. In A. A. Wager, & D. W. Stinson (Eds.), *Teaching mathematics for social justice: Conversations with mathematics educators* (pp. 49-62). National Council of Mathematics Teachers.

Geiger, V., Beswick, K., Fielding, J., Scheiner, T., Kaiser, G., Goos, M., & Fernandez, K. (2023). Investigating critical mathematical thinking when applying mathematics to real-world problems. In *Thirteenth Congress of the European Society for Research in Mathematics Education* (pp. 1177-1184). Alfréd Rényi Institute of Mathematics and Eötvös Loránd University Budapest, Hungary.

González, O., & Chitmun, S. (2017). Using sport-related socially open-ended problems to coordinate decision-making, variability modeling and value awareness. In T. Dooley, & G. Gueudet (Eds.), *Proceedings of the tenth congress of the European society for research in mathematics education* (pp. 780-787). DCU Institute of Education and ERME.

Hattori, Y. (2017). Fostering critical thinking in mathematics education and its challenges. In *Proceedings of the 5th Spring Research Conference* (pp. 269-276). Japan Society of Mathematical Education.

Hattori, Y. (2020). Characteristics of critical thinking induced by socially open-ended problem: Comparison of elementary and junior high school students in the same problem. In *Proceedings of the 8th Spring Research Conference* (pp. 161-168). Japan Society of Mathematical Education.

Hattori, Y., & Fukuda, H. (2019). Aspects of fair-minded critical thinking based on the perspective of critical mathematics education: A case of lesson practice in the lower secondary education. In *Proceedings of the 7th Spring Research Conference* (pp. 19-26). Japan Society of Mathematical Education.

Hattori, Y., & Inoue, Y. (2015). The development of RLA in teaching mathematics to foster critical thinking: Peer reviews by students. *Journal of JASME Research in Mathematics Education, 21*(2), 1-12. https://doi.org/10.24529/jasme.21.2_1

Hattori, Y., Inoue, Y., Matsubara, K., Hakamata, R., & Hisadomi, Y. (2023). Development and practice of teaching materials in high-school mathematics with a focus on the nurturing and evaluation of critical thinking: A socially open-ended problem"Propose a Mayonnaise Nozzle". *Journal of JASME Research in Mathematics Education, 28*(2), 77-97. https://doi.org/10.24529/jasme.28.2_77

Hattori, Y., Fukuda, H. & Baba, T. (2021). Development of socio-critically open-ended problems for critical mathematical literacy: A Japanese case. *Journal of Educational Research in Mathematics, 31*(3), 357-378. https://doi.org/10.29275/jerm.2021.31.3.357

Hisadomi, Y., & Koyama, M. (2018). A practical research on incorporation of performance assessment in the end of teaching unit in high school mathematics: Through analyzing the qualitative assessment of reports with the rubrics by students and teachers. *Journal of JASME Research in Mathematics Education, 24*(2), 37-49. https://doi.org/10.24529/jasme.24.2_37

Ito, T. (2015). A fundamental research on fostering of the critical thinking in the elementary mathematics education: Based on counterexample. *Journal of JASME Research in Mathematics Education, 21*(2), 39-48. https://doi.org/10.24529/jasme.21.2_39

Jablonka E. (2020) Critical thinking in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education*. Springer, Cham. https://doi.org/10.1007/978-3-030-15789-0_35

Kacerja, S., & Julie, C. (2023). Values in preservice mathematics teachers' discussions of the body mass index: A critical perspective. *The Journal of Mathematical Behavior*, *70*, Article 101035. https://doi.org/10.1016/j.jmathb.2023.101035

Kincheloe, J. L., McLaren, P., Steinberg, S. R., & Monzó, L. (2017). Critical pedagogy and qualitative research: Advancing the bricolage. In N. K. Denzin, & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (5th ed., pp. 235-260). Sage.

Krulik, S., & Rudnick, J. A. (1999). Innovative taskes to improve critical - and creative - thinking skills. In L. V. Stiff, & F. R. Curcio (Eds.), *Developing mathematical reasoning in grades K-12* (pp. 138-145). Reston, VA: National Council of Teachers of Mathematic.

Kubo, Y. (2016). A study on critical thinking in mathematics education. *Proceedings of the 4th Spring Research Conference,* 97-104. Saitama, Japan: Japan Society of Mathematical Education.

Kubo, Y. (2019). University students' attitudes on relevance in critical thinking. In *Proceedings of the 7th Spring Research Conference* (pp. 27-34). Japan Society of Mathematical Education.

Michita, Y. (2003). Diversity and a fundamental image of the major concepts of critical thinking. *Japanese psychological review, 46*(4), 617-639.

Michita, Y. (2005). Some notes on the critical thinking in a strong sense, bulletin of college of education. *University of the Ryukyus, 66*, 75-91.

Michita, Y. (2013). Mittu no toi kara hihantekisikouryokuikusei ni tuite kangaeru (Thinking about the development of critical thinking through three questions). *Psychology World, 61*, 9-12.

Morioka, K., Shiobara, T., & Honma, Y. (1993). *Sin syakaigakujiten* (New sociological dictionary). Yuhikaku Publishing.

Organisation for Economic Co-operation and Development. (2018). *PISA 2021 mathematics framework* (Second draft). https://pisa2021-maths.oecd.org/files/PISA 2021 Mathematics Framework Draft.pdf

Paul, R. W. (1992). Critical thinking: What, why, and how. *New Directions for Community College, 77*, 3-24. https://doi.org/10.1002/cc.36819927703

Paul, R. W. (1995). *Critical thinking: How to prepare students for a rapidly changing world.* Santa Rosa, CA: Foundation for Critical Thinking.

Sagiv, L., Roccas, S., Cieciuch, J., & Schwartz, S. H. (2017). Personal values in human life. *Nature Human Behaviour*, *1*(9), 630-639. https://doi.org/10.1038/s41562-017-0185-3

Shimada, I. (2016). Possibility of developing critical thinking through socially open-ended problems. In *Proceedings of the 4th Spring Research Conference* (pp. 113-120). Japan Society of Mathematical Education.

Shimada, I. (2017). *Sansu sugaku kyouiku to tayou na kachikan: Syakaiteki open-end na mondai ni yoru torikumi* (Arithmetic and mathematics education and diverse values: Approaches through socially open-ended problems). Toyokan Publishing.

Skovsmose, O. (1994). *Towards a philosophy of critical mathematics education*. Springer. https://doi.org/10.1007/978-94-017-3556-8

Skovsmose, O. (2016). What could critical mathematics education mean for different groups of students? *For the Learning of Mathematics*, *36*(1), 2-7.

Skovsmose, O. (2020). Critical mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education.* Springer, Cham. https://doi.org/10.1007/978-3-030-15789-0_34

Skovsmose, O. (2022). Concerns of critical mathematics education-and of ethnomathematics. *Revista Colombiana de Educación*, (86), 365-382. https://doi.org/10.17227/rce.num86-13713

Skovsmose, O., & Nielsen, L. (1996). Critical mathematics education. *International Handbook of Mathematics Education* (part two) (pp. 1257-1288). Kluwer Academic Publishers.

Vincent-Lancrin, S., González-Sancho, C., Bouckaert, M., de Luca, F., Fernández-Barrerra, M., Jacotin, G., Urgel, J., & Vidal, Q. (2019). *Fostering students' creativity and critical thinking: What it means in school, educational research and innovation*. OECD Publishing, Paris. https://doi.org/10.1787/62212c37-en