

RESEARCH

Open Access



Application focused on structural comprehension of mathematics contextual problems for kindergarten students

Pedro Gabriel Fonteles Furtado^{1*} , Tsukasa Hirashima¹, Yusuke Hayashi¹ and Kazushige Maeda²

*Correspondence:

gabrielpff2@gmail.com

¹Graduate School of Engineering,
Hiroshima University, 1-4-1
Kagamiyama, Higashi-Hiroshima
739-8527, Japan

Full list of author information is
available at the end of the article

Abstract

There has been a lack of research into conceptually teaching the structure of contextual problems to kindergarten students. This is despite the popularity of contextual problems and the difficulties that students have with them. This study focuses on creating a textless, image-based application that uses the Triplet Structure Model to teach about the structure of contextual problems. This model has been successfully used in the past to teach Japanese elementary school students. The pilot study suggests that young children are successful at learning about the structure of contextual problems by using the application. The results suggest that users' interactions are not thoughtless and that the interactions improve as they use the application.

Keywords: Word problems, Kindergarten, Problem-based learning, Triplet Structure Model, Contextual problems

Introduction

Contextual problems play a major role in mathematical education. Contextual problems are defined as problems with experientially real contexts in the Realistic Mathematics Education approach (Gravemeijer and Doorman 1999). These problems can function as the basis for connecting informal and formal knowledge of mathematics. Because of this, they can help students better grasp formal mathematics.

Students usually encounter contextual problems in early arithmetic classes in their first year of elementary school. These contextual problems describe an addition or a subtraction. Students are often required to:

- 1 Write the arithmetic expression corresponding to a certain contextual problem.
- 2 Calculate the answer by solving the arithmetic expression.

Focusing on the first step, the work of Rivera (2014) states that elementary contextual problems need to be conceptually understood, adding that teachers often teach children to rely on keywords to transition to the arithmetic expression. The author also states that this keyword-based method shows a lack of conceptual understanding that is necessary for solving harder problems. An analysis by Hegarty et al. (1995) also points out this issue, stating that students that rely on keywords instead of creating a mental model of the problem are usually unsuccessful.

The work of Greer (1997) describes a similar problem, stating that students have a disregard for the reality described in the problems, relying instead on certain words in the text to select an operation and extracting numbers, without modeling the situation in their mind.

One way to handle this situation is by doing interventions before the children start elementary school. Proficiency in mathematics in the early years has been shown to predict performance in future years (Jordan 2010). Moreover, interventions in kindergarten, such as in the work of Griffin (2004), have been effective in improving the math skills of young children. As such, an intervention on kindergarten focused on creating mental models of contextual problems should be effective in preparing the children for the years to come.

Based on the above, the following question is posed: How to help kindergarten students develop this conceptual understanding that allows for mentally modeling contextual problems?

One proposed solution is the use of the Triplet Structure Model. The Triplet Structure Model can help students to model early mathematical problems by creating a three-sentence representation bridging the contextual problem and its arithmetic expression, as shown by Hirashima et al. (2014). This model will be further explained in later sessions. Monsakun is an interactive learning environment for learning by problem-posing of arithmetic contextual problems, and it is designed based on the Triplet Structure Model. In Monsakun, a learner poses an arithmetic contextual problem by arranging three sentences from a group of cards. Monsakun diagnoses the posed problem based on the Triplet Structure Model and gives feedback. It has been used in elementary schools and has shown promising results thus far, with experiments suggesting that the use of Monsakun results in better problem-solving skills. Examples of this include forward-thinking addition and subtraction problems (Hirashima et al. 2007), reverse thinking addition and subtraction problems (Hirashima and Kurayama 2011; Yamamoto et al. 2012), and multiplication and division problems (Yamamoto et al. 2014). The problem-posing process has also been analyzed for addition and subtraction problems (Hasanah et al. 2017; Supianto et al. 2017). Results suggest that the Triplet Structure Model can be used in the classroom to deepen conceptual understanding of contextual problems.

However, Monsakun is not suitable for kindergarten students, since not all of them are accustomed to reading. As far as we know, there are no tools that focus on the conceptual understanding of contextual problems that do not rely on reading. As such, presently, there are no such tools that can be used by kindergarten students. The present study presents an application design based on the Triplet Structure Model that uses speech synthesis and images. This study seeks to verify if kindergarten students can interact with the conceptual structure of contextual problems through the Triplet Structure Model. This is an ongoing effort on conceptually teaching contextual problems to young, Japanese children (Furtado et al. 2017). Our research questions are (1) whether kindergarten students can meaningfully interact with the Triplet Structure Model and (2) whether the interactions of the students with the Triplet Structure Model improve as the students use the application.

Background

Past research has shown that proficiency in mathematics in the early years, such as number sense, can predict mathematical performance in later school life (Jordan 2010). This

has been verified all the way to the third year of elementary school. It suggests that building mathematical skills before entering elementary school can greatly impact how the student deals with mathematics in the coming years. Hence, interventions at an early age could be successful in improving students' understanding of contextual problems.

Interventions to strengthen mathematical skills for both kindergarten and pre-kindergarten students have been implemented previously. The work of Griffin (2004) focuses on cognitive developmental theory to define number sense and teach it through a series of activities and games. It has found success in increasing students' understanding of number sense. The intervention of Dyson et al. (2013) also saw success, with activities based on various mathematics curricula designed for kindergarten students. The work by Starkey et al. (2004) focused on pre-kindergarten children of around four years of age. Their study focused on more than just number sense, also considering subjects such as geometric reasoning, and their results were positive.

Computer-based interventions have also been researched. The work by Wilson et al. (2009) used adaptive software that models the progress of the children and regulates problem difficulty for each child, resulting in software that can be used without supervision. The software is focused more on specific skills and on fast and accurate access, so the authors state that what is developed in the software is number sense access instead of number sense itself, describing the software as a complementary activity for children who are lagging behind in class. Clements and Sarama (2007) used a software called "building blocks." Building blocks focus on many aspects of math teaching, including number sense, and targets children from pre-kindergarten to grade 2. This method has found good success.

These interventions, in general, focus on whole number understanding. However, they do not focus on the conceptual understanding of contextual problems, which is a big part of how students interact with mathematics after kindergarten.

Computer-assisted interventions focused on contextual problems

With regard to contextual problems, interventions using computers have been researched before. The work of (Kim: Computer-assisted Mathematical Interventions on Word Problems for Elementary Students with Underachievement in Mathematics, unpublished) provides a review of the literature on the use of computer-assisted interventions (CAI) focused on contextual problems for children with math difficulty. The review suggests that CAI are useful for supporting students with math difficulty. It also suggests that CAI can be more effective than paper-based approaches. However, the review suggests that there is not enough research to answer if computer-mediated learning is more effective than teacher-mediated learning.

The study conducted by Freiman et al. (2017) approached the problem by using algebraic thinking in the early grades. The study argued that analyzing the problems through this type of thinking required a structural understanding of the problems. The analysis of how students use the software shows that students were gradually shifting to an algebraic way of thinking while solving and understanding the problems. However, the tool deployed elements that would be hard to adapt for kindergarten students, such as text and diagrams.

The work of Yerushalmy (2006) focuses on a tool based on graphs to help students solve algebra contextual problems. The tool is mainly used to plot graphics, to give more visual

context to the equations, and to approximate values. The author states that it would be useful for the tools to become a bridge between the contextual problems and the algebraic symbols. While targeting contextual understanding, an approach using graphics would be too complex for young children. In the study by Koedinger and Anderson (1998), which also considered algebraic problems, a cognitive tutor attempted to model students and provide specific tasks for each of them. The study analyzed the problem-solving strategy variants of a computerized tutor and explored the differences between the strategies. The cognitive tutor focused more on guiding students in the various steps of problem-solving instead of focusing on understanding the problem itself. The work of Hamadneh et al. (2016) had elemental school students with disabilities interact with math content through PowerPoint. Users would revise the content and solve problems while receiving feedback through the system. The score of system users was significantly higher than the control group. The test used for evaluation included addition, subtraction, multiplication, and division. While the content of the system has not been detailed, it is implied that students mostly solved word problems and received feedback on the correctness of their answer. The system used relies on reading and might not have a step focused on contextual understanding.

The work of Huang et al. (2012) describes a problem-solving software based on dividing the problem into phases and assisting the students in solving each phase. One of the phases is dedicated to conceptually understanding the problem and another to planning and carrying out the calculations. Experimental tests were also performed. The experimental group showed better scores than the control group and participants showed a positive affective response towards the system. While the system deployed does have a step focused on conceptual understanding of the problems, kindergarten students who are unable to read would not be able to use it.

The above research examples show various degrees of success in using computers to help students with contextual problems. Some of the work even have a focus on structural understanding. However, all of them rely on text and on procedures that would be too complicated for kindergarten students.

The main contribution of our study is that the system we propose focuses on the conceptual understanding of contextual problems while being usable by kindergarten students. Furthermore, we have analyzed collected data from the software to investigate whether students are engaged in thinking about the structure of contextual problems. Whether or not there were improvements in the conceptual understanding of the students is also investigated. Such a study, as far as we know, has not been done before.

Design and development of the application

The development process

The development team consisted of four people. All members influenced the design of the application. One member of the team was in charge of programming, designing the screens, and creating the graphics. One of the team members had years of experience as a primary school teacher. Given his understanding of how to handle children and of teaching mathematics, he was able to point out various shortcomings in the application. He also had experience working with Monsakun inside the classroom. The other two members contributed by helping design the activities and by suggesting possible solutions to the problems that were pointed out by the team.

Developing the software involved creating features, having the team members test and analyze the current state of the software, and then deciding on improvements, additions, and cuts. As such, the development process was iterative in nature (Larman and Basili 2003). The team would have a meeting and decide on the tasks that had to be completed until the next meeting. The time interval between the meetings was defined based on estimations made by the developer. This process continued until all four team members were satisfied with the application.

Ideally, the constructed activities should have been tested with children during development. However, the tight schedule of kindergarten in Japan has made this unfeasible. The presence of a person who has experience as a primary school teacher is thought to mitigate this issue, but not to eliminate it. The use of pictures similar to Japanese textbooks in the application and software demonstrations with kindergarten teachers also helped mitigate this issue.

Triplet structure model

The Triplet Structure Model binds the three quantities of one arithmetical operation (operand, operant, and result quantities) to contextual story roles by using three quantity sentences. This is illustrated on Fig. 1. It is usually used to describe arithmetic contextual problems where one of the quantities is unknown.

These quantity sentences can be classified into two types: independent quantity and relative quantity sentences. Independent quantity sentences state the existence of a certain number of objects. For example, “there are two apples.” Relative quantity sentences depend on the previous existence of a certain number of objects. “Two apples were eaten” is an example of relative quantity. Relative quantities operate on one or more independent quantities.

The example in Fig. 1 contains a case of subtraction story. The Triplet Structure Model refers to this type of story as a decrease story. An example of an increase story is given below:

- 1 There are five apples (independent quantity).
- 2 Two apples were brought (relative quantity).
- 3 There are seven apples (independent quantity).

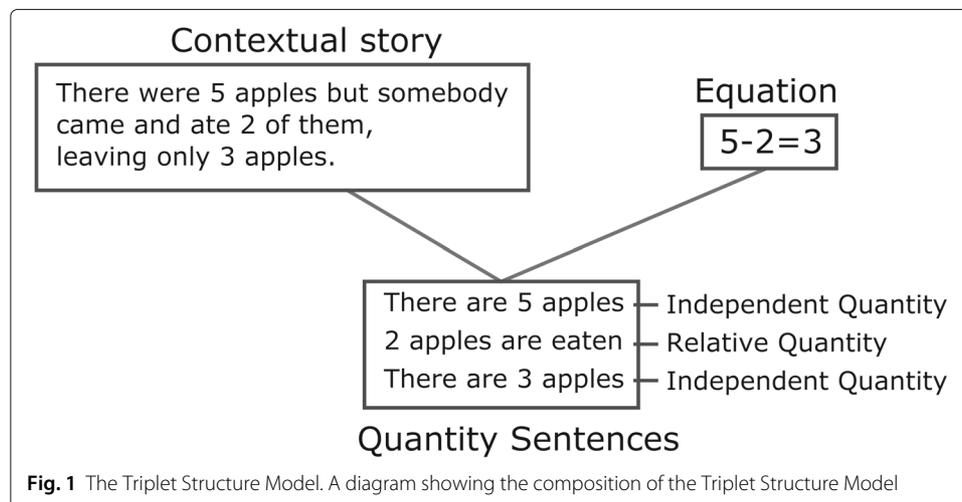


Fig. 1 The Triplet Structure Model. A diagram showing the composition of the Triplet Structure Model

And for combination stories:

- 1 There are three apples (independent quantity).
- 2 There are two oranges (independent quantity).
- 3 Put together, there are five apples and oranges (relative quantity).

Both increase and decrease stories are composed of one independent quantity in the beginning, one relative quantity in the middle, and another independent quantity in the end. The relative quantity shows how much the amount of the object changed (increased or decreased, depending on the story), while the independent quantities show the number of objects before and after the change. All three quantities in increase and decrease stories must refer to the same object, or else the story is invalid (“there are two apples, one apple is eaten, there is one banana” is not valid, for example).

Composition stories have a slightly different structure. In composition stories, the relative quantity comes at the end, with two independent quantities coming before it. The two independent quantities must then refer to different types of objects, while the relative quantity describes the total number of the two objects together.

In decrease and increase stories, the relative quantity created a change in the amount of a certain object. In the composition case, the relative quantity provides a numeric observation of the previously defined quantities without changing their value. Different objects here could be “John’s apples” and “Mary’s apples,” while the relative quantity would be “John and Mary’s apples put together.”

There is one more story type that is outside of the scope of this paper, namely, the comparison story. Due to the nature of this story, we found it difficult to show this by using pictures and decided to not include it in our design. However, it may be included in future research after a satisfactory representation is found. More information on this and on the Triplet Structure Model can be found in the work of Hirashima et al. (2014).

Application design

The application design must allow for interactivity with the Triplet Structure Model without relying on text. Our solution has been to use pictures and spoken sound. There are two types of pictures. The first type is rectangular in shape and is called an overall story picture. This type shows the entire problem at once. The second type is the story piece picture. These are small and square and represent each sentence in the Triplet Structure Model. Understanding the relationship between the overall story pictures and the story piece pictures is like understanding how a problem is structured in the Triplet Structure Model. The design of the pictures will be further introduced below. In the application, we also ask for users to connect the pictures to numbers. This connection brings them a step further to connect the in-context parts of a problem to the out-of-context parts of an arithmetic expression. This type of connection is in-line with the conceptual understanding of contextual problems described in previous sections.

We have divided the application activities into levels. Each level is described in more detail below. Level 3 is critical to the application. It focuses on connecting story piece pictures to big pictures. Performing this requires students to be able to divide the big picture into three small parts, each related to a meaningful quantity.

Level 6, which connects the pictures to numbers, is also worth noting. First, we ask students to connect story piece pictures to numbers. Since each small picture refers to

one quantity, connecting one meaningful quantity to one number is not a difficult task. However, later, the application requests users to connect one big picture to three numbers. Since no other help is given to the user, they have no choice but to visualize the three numbers related to the big picture. This is like writing the arithmetic expression of a given problem while being given only the picture of the problem. Students that can perform this task well should be able to meaningfully connect contextual problems to their arithmetic expressions. These levels will be explained further below.

Image and sound design in connection with the Triplet Structure Model

The two picture representations of a Triplet Structure Model story can be seen in Fig. 2. Each picture also has an accompanying sound. On the left, the overall story is shown, where one picture contains all the information to describe a story. On the right, story pieces are shown, in which three pictures are put together to describe a story.

While in the model, we would have “There are three apples,” and in the picture’s accompanying sound, we would have “there are three apples on the shelf.” We describe the place that the objects are in to strengthen the connection between what the sound is saying and what the picture is showing. Story piece pictures come in two types, independent and relative pictures. They correspond to independent and relative quantities in the Triplet Structure Model. Independent pictures are usually composed of stationary objects. Relative pictures will usually describe an action, like a boy inserting objects somewhere or an animal entering a place. By describing an action or movement, we can convey the same idea as the relative quantities of the Triplet Structure Model. We can be confident in the children’s interpretations of the pictures because very similar designs are used in the textbooks used in Japanese schools.

Overall story pictures represent the entire problem in a single picture. In the overall story picture in Fig. 2, the three numbers of the problem can be seen as:

- 1 The number of watering cans on the shelf.
- 2 The number of cans the boy puts on the shelf.
- 3 The total number of cans after the boy places them on the shelf.

The three numbers are mapped to the three-story piece pictures, creating the relationship between the overall picture and the story pictures. The sound related to the overall picture is a simple combination of the sounds of the three-story pieces put together. While connecting the overall story to the three pictures may seem like a trivial task, it is not so simple. While the first two numbers are quite clear in the overall story picture, the third number, which represents how many cans there will be on the shelf, requires more thinking. Students have to understand the described story and then recognize that there will be two watering cans on the shelf after the boy has finished. This number could be calculated

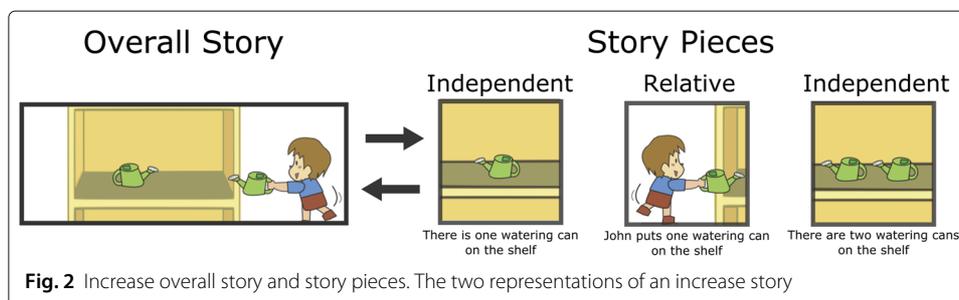


Fig. 2 Increase overall story and story pieces. The two representations of an increase story

by counting or by mental addition, either is valid. What matters is if the student can interpret the story or not. Since not all quantities are explicitly shown, connecting the overall story picture to the three-story piece pictures requires more than simply looking at the photo and trying to match the objects or scenery.

Level 1

At level 1, participants listen to audio describing a picture and then choose, from three pictures, which picture corresponds to the audio. At this level, the pictures are story pieces and not the overall story pictures. This is an introductory stage to introduce the pictures that can make up a story and their corresponding description.

Figure 3 shows an example. In this case, the student would hear one of these three phrases spoken out loud:

- 1 There are two watering cans on the shelf.
- 2 There is one watering can on the shelf.
- 3 John puts a watering can on the shelf.

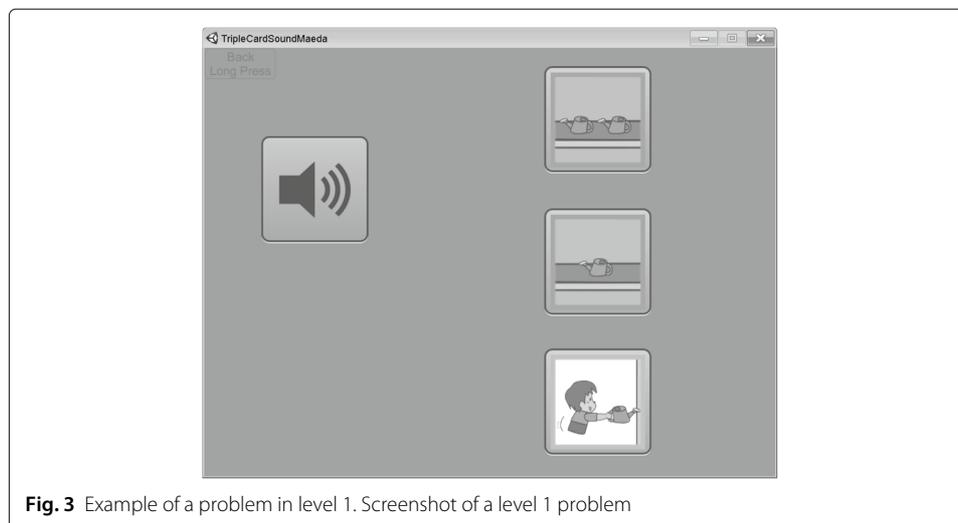
Then, the student would have to touch the picture that corresponds to the phrase that they heard. The student can also repeat the sound by pressing the sound button.

Level 2

Level 2 focuses on connecting the overall story pictures to their spoken narration, to ease students into understanding the content of the pictures. Unlike story piece pictures, the pictures in this level have a description comprised of three phrases, each one corresponding to one of the quantities described in the story.

Level 2 is made up of two parts, level 2-1 and level 2-2.

Level 2-1 This level is based on true or false problems. Participants hear the spoken narration and are shown one overall story picture. They must decide if the picture described in the narration is the picture being shown or not by pressing true or false.



Level 2-2 Part two is like level 1, where participants are shown three pictures and audio, having to decide which picture corresponds to the audio.

Level 3

Level 3 focuses on connecting overall stories to their story piece pictures, and it is made up of two parts.

Level 3-1 In this part, we have true or false problems, with participants being shown one overall story big picture and three-story piece pictures. They then must decide if the three-story piece pictures correspond to the same story being shown in the big picture or not by choosing from true or false buttons.

Level 3-2 Participants are shown an overall story picture and given multiple story piece pictures in this part. They are asked to use three of the story piece pictures to make up a single story. The made-up story must correspond to the same story being shown in the overall story picture. Figure 4 illustrates this setup. Participants are given five pictures. Since only three pictures make up a story, the remaining two pictures are dummy pictures. Dummy pictures are added to give students more to think about. Problems in Level 3-2 do not all have the same difficulty, the number of blank spaces varies as described below:

- 1 One blank: Problems 1 to 5. There is only one blank space. The other two spaces are automatically filled. The player can move three of five cards.
- 2 Two blanks: Problems 6 and 7. Two blank spaces. One space is automatically filled. The player can move four of five cards.
- 3 Three blanks: Problems 8 to 11. No blank space is filled. The player can move all five cards. This difficulty progression is used to ease the participants into working with level 3 and constructing a story from three pieces. We stress once again that this is a key skill in the context of the Triplet Structure Model.

Level 4

Level 4 requires participants to listen to a narration of a story, and then, they are tasked with forming the story by using story pieces. This is similar to the second part of level 3.

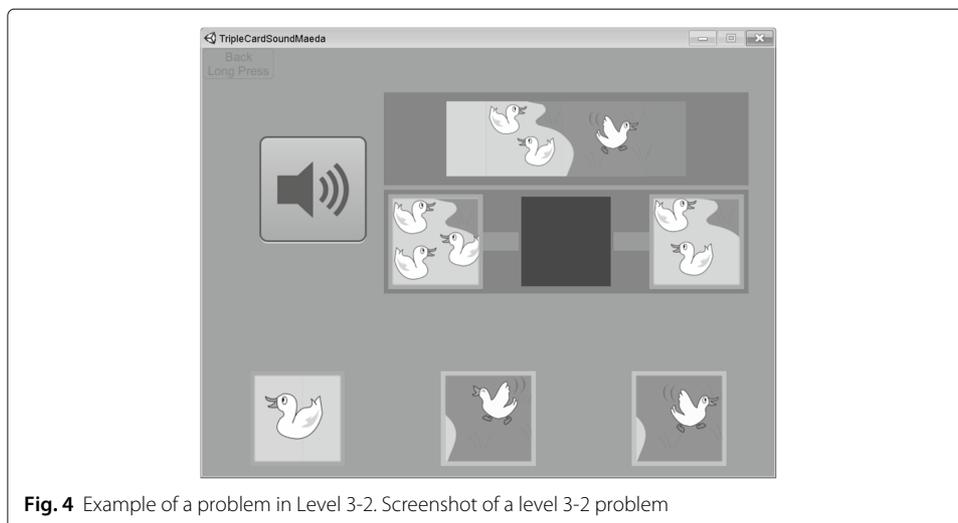


Fig. 4 Example of a problem in Level 3-2. Screenshot of a level 3-2 problem

The difference is that in level 3, the correspondence was with the overall story picture, while in level 4, it is with the spoken narration. This level is composed of eight problems, with the first three being easier, only allowing users to move three pictures out of five. The rest of the problems allow the user to move all five pictures.

Level 5

Level 5 shows participants an overall story picture and asks if that story belongs to a certain story type (the types being “increase,” “decrease,” and “combine”), with the participant having to choose “true” or “false.” This level relates to how well the participant understands the concepts of increasing, decreasing, and combining. It also relates to how the participants comprehend the story of each picture.

Level 6

In this level, participants connect numbers to pictures. It is divided into two parts. In part one, students are tasked with connecting numbers to story piece pictures. In part two, students are asked to connect three numbers to a single overall story picture.

The setup that can be seen in Fig. 5. As stated before, this is a problem that requires a deep understanding of the three quantities that can be interpreted from a single picture. Users that can do this should be able to construct mental models that allow them to be successful problem solvers.

This section introduced the application’s design, sound design, picture design, and explained each level of the application. Which levels are critical to the conceptual understanding of contextual problems and why they are critical has also been discussed.

The pilot study

One hundred Japanese kindergarten students participated in the experiment. They were around five years old, and they were divided into three classes. All students belonged to the same kindergarten. The application was introduced in around 10 min. During this time, four to five problems of the first two levels were shown by using a projector connected to a tablet. The participants were encouraged to give their opinion on the answer while the teacher advanced through the problems. Afterward, students had around 20 min

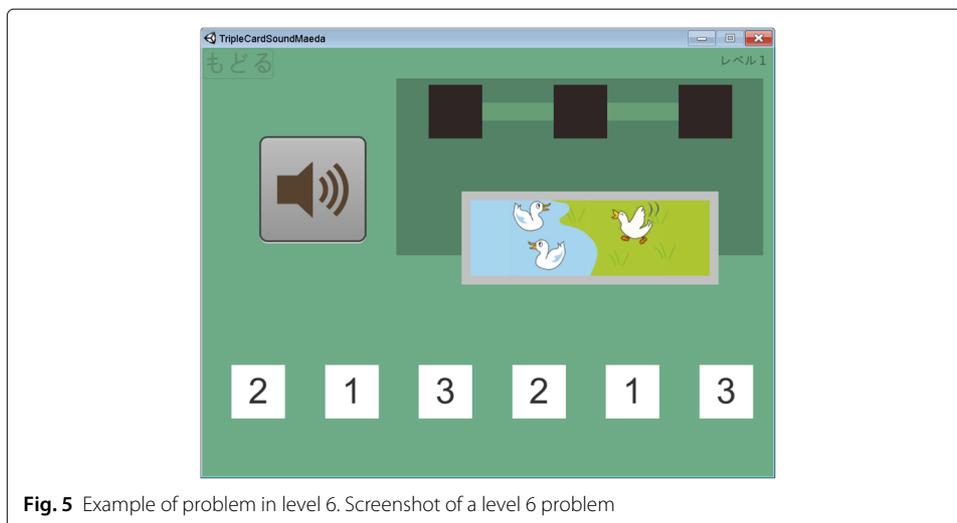


Fig. 5 Example of problem in level 6. Screenshot of a level 6 problem

to interact with the application by themselves, by using an android tablet that contained the application and headphones. There was one tablet available per student, so no sharing was necessary. The use of the application by the students and the explanations were recorded on video.

Afterward, user log data was extracted from the tablets for analysis. Because of privacy issues, it was not possible to associate student information recorded on video to their respective log data. The data includes various events that can happen during application use. Examples of these events are a list of problems being started or a button being pressed. A snippet of data can be seen below.

```
config choice , LevelConfig , 2016-12-31-12:26:26, 11.6 ,
level choice , 0, 2016-12-31-12:26:30, 15.16 ,
problem list start , ProblemList - Level 1 ,
                                2016-12-31-12:26:30, 15.16 ,
problem in list start , 0, unitCardVsSound ,
                                2016-12-31-12:26:30, 15.18 ,
problem in list start , 1, unitCardVsSound ,
                                2016-12-31-12:26:32, 16.84 ,
problem solved , 2016-12-31-12:26:32, 16.84 ,
problem mistake , jouro1boyadd , 2016-12-31-12:26:32, 17.66 ,
```

This raw data was then processed into an intermediary comma separated value (CSV) format for easier analysis. The main interest of this intermediary data was the number of mistakes that each user made while trying to solve the problems. An example of this data can be seen in Table 1.

The measured performance of participants was compared to a calculated “gaming the system” performance metric. Gaming the system (GTS, will also be used as “to game the system”), refers to the behavior displayed when students attempt to systematically take advantage of the way the system is made. One way users can GTS is by randomly picking options. This definition and further discussion can be found in the work of Baker et al. (2008). We used the probability that students will solve a level on their first try as a performance metric. This probability depends on both the level and the difficulty of the problem.

For example, on level 4, the probability, by random chance, of solving the first three easy problems in one try is 1/3 (only one option with three choices). The probability of solving the last five problems is 1/60 (five choices on the first card, four choices on the second card, and three choices on the third card). The calculation $(3 \times 1/3 + 5 \times 1/60)/8$ gives us the probability of solving a problem in level 4 in one try, by random chance. We can use a similar logic to calculate the average number of attempts.

Table 1 Example of intermediary CSV generated from raw data

Session	List	Problem	Mistakes	Operation	Difficulty	Probability
5	ProblemList level 1	0	2	Exist	1	0.3333333
5	ProblemList level 1	1	0	Add	1	0.3333333
5	ProblemList level 1	2	1	Exist	1	0.3333333
5	ProblemList level 1	3	1	Exist	1	0.3333333

However, the calculated values are based on which problems the students have attempted on each level. While uncommon, students could have stopped midway through a level, and restarted and completed the level. This means that the calculated values are based on which problems the participants have completed for this study.

To analyze the data collected from the students, we used two metrics. The first was the average number of attempts per problem. That is, the average number of times they try to get to the correct answer. The second metric was the ratio of problems solved in one try to the total number of problems attempted (RPOT). The metrics are calculated for every student/level pair. The first metric showed how much difficulty a participant had with the problem. The second one revealed signs of problem mastery, since solving the problem in one try implies that the student obtained the required knowledge related to the problem. We also analyzed peaks of difficulties. That is, we analyzed when students are first introduced to harder problems and how they coped with the problem when they saw it again, by comparing their performances.

For example, looking at the data in Table 1, the student related to session 5 took three attempts to solve problem 1, one attempt to solve problem 2, and two attempts to solve problems 3 and 4. If these four problems were the only problems in level 1, the average for this student for level 1 would be two trials per problem. Since he only solved one problem in one try, his RPOT would be 0.25.

Statistical analyses were run in the group of students as a whole using the metrics defined above. Furthermore, the interactions of students that performed similarly to the calculated GTS metrics were individually analyzed as use cases.

A pre-test and a post-test were not included in our pilot study due to constraints on time in the schedule of the school. Log data was only collected when the user completed a level. This means that our analysis did not include some data from levels that were stopped midway through.

Results and discussion

Four thousand six hundred sixty-five lines of raw data were collected. To prepare for the analysis, the data was converted into 2578 lines of intermediary CSV. No data was eliminated.

Overall analysis

Students were quick to progress through levels 1 and 2 and found difficulty with level 3. During application use, students were deeply focused on the application. Other than using the software, they commented on the progress of each other and asked for help from the teachers and helpers. This has been interpreted as a positive affective response to the software. The number of users who completed each level can be seen in Table 2.

The average number of attempts per problem and the RPOT calculated for all levels were calculated by using the collected data. Likewise, gaming the system (GTS) versions of the same metrics were calculated. Tables 3 and 4 show that the difference between student's performance and the calculated GTS performance is statically significant for both measures. This result shows that the strategies students employed to solve each level were more effective than GTS. This suggests that students were not blindly progressing through the application without thinking. Furthermore, it is indicative that students displayed the necessary skills to solve each level. This addresses our first research question

Table 2 Number of participants who cleared each level of the application

Level	N
1	100
2-1	90
2-2	89
3-1	20
3-2	14
4	13
5	9
6	2

that kindergarten students are able to meaningfully interact with the Triplet Structure Model.

It also shows signs that our pictures, and their audio descriptions, fit with the participants’ interpretations of the pictures, otherwise, they would not have been able to perform well on the first two levels. Results suggest that students were able to use the software and their observed affective response was also positive.

Between level 2-2 and 3-1 there, was a large drop from 89 students to 20 students. This drop has been accounted to time constraints. Given how level 3–1 contained only truth-and-false exercises, there is little reason to believe that students would not be able to complete it given more time.

Level 3-2 Figure 6 shows the average number of attempts of each problem during GTS for level 3-2. Students had little trouble with difficulty 1. The same cannot be said for difficulties 2 and 3. There are two spikes in the graph which correspond to the difficulty changes. This shows that difficulty 1 was not so demanding, and as a result, the students were not pushed to understand the meaning of what they were doing. As soon as the difficulty went up, students struggled to solve the problem. However, after this first peak, their performance quickly increased. This means that students were quick to understand what was being asked of them after solving it for the first time.

Figure 7 shows a similar trend. The figure shows the value of RPOT for level 3-2.

Students had little trouble with difficulty 1, and they had difficulties with the same problems shown in the previous analysis. However, it is worth noting that around 20% of the participants managed to solve problem 8 in one try. Level 8 is the point at which difficulty 3 is first introduced. This means that a good portion of the students had a good understanding before the difficulty increased. The effects of the peaks are summarized

Table 3 One sample *t* test results comparing measured number of attempts per problem to the equivalent “gaming the system” calculated value. sd stands for standard deviation

Level	Avg. no. of attempts	Avg. no. of attempts (GTS)	<i>t</i>	df	<i>p</i>
1	1.16 (0.24)	2.00	−34.79	99.00	< 0.001
2-1	1.17 (0.23)	1.50	−13.61	89.00	< 0.001
2-2	1.24 (0.25)	2.00	−29.31	88.00	< 0.001
3-1	1.30 (0.32)	1.50	−2.77	19.00	0.01
3-2	4.77 (3.80)	13.18	−8.28	13.00	< 0.001
4	1.44 (0.66)	18.44	−93.11	12.00	< 0.001
5	1.19 (0.18)	1.50	−5.38	8.00	< 0.001
6	1.31 (0.27)	8.25	−37.00	1.00	0.02

Table 4 One sample *t* test results comparing measured ratio of problems solved in 1 try to the equivalent “gaming the system” calculated value

Level	Ratio of solved in 1 try	Ratio of solved in 1 try (GTS)	<i>t</i>	df	<i>p</i>
1	0.87 (sd = 0.16)	0.33	33.13	99.00	< 0.001
2-1	0.84 (sd = 0.18)	0.50	18.55	89.00	< 0.001
2-2	0.82 (sd = 0.18)	0.33	26.11	88.00	< 0.001
3-1	0.75 (sd = 0.27)	0.50	3.99	18.00	0.001
3-2	0.56 (sd = 0.27)	0.21	4.82	13.00	< 0.001
4	0.81 (sd = 0.24)	0.17	9.43	12.00	< 0.001
5	0.81 (sd = 0.18)	0.50	5.38	8.00	0.001
6	0.88 (sd = 0.00)	0.10	-	-	-

Data was constant for level 6, so the test was not performed

in Fig. 8. The only peak that shows a statistical significance ($t(13.83) = 2.21, p < .05$) is the difference between the average number of attempts between the first problem ($M = 18.07, SD = 22.35$) and the other problems ($M = 4.67, SD = 6.89$). Although the other peaks that were analyzed did not show statistical significance, they all show the same trend.

These results show that students had trouble when the difficulty goes up in level 3-2 but, after initially struggling, they could understand what was being asked of them, and showed patterns of growth.

While this trend has been verified for the group in general, it remained unknown if the students who had most trouble also showed similar patterns. Some of these students have shown performance similar to the predicted GTS metrics. To investigate whether these students also took advantage of using the system, we perform case studies for every student whose average number of tries went above 50% of the predicted values.

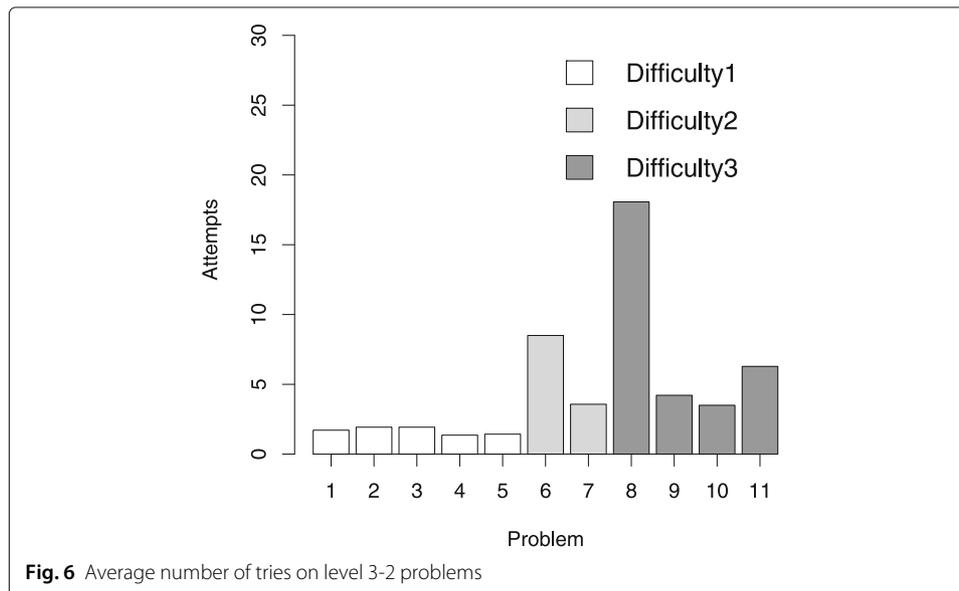


Fig. 6 Average number of tries on level 3-2 problems

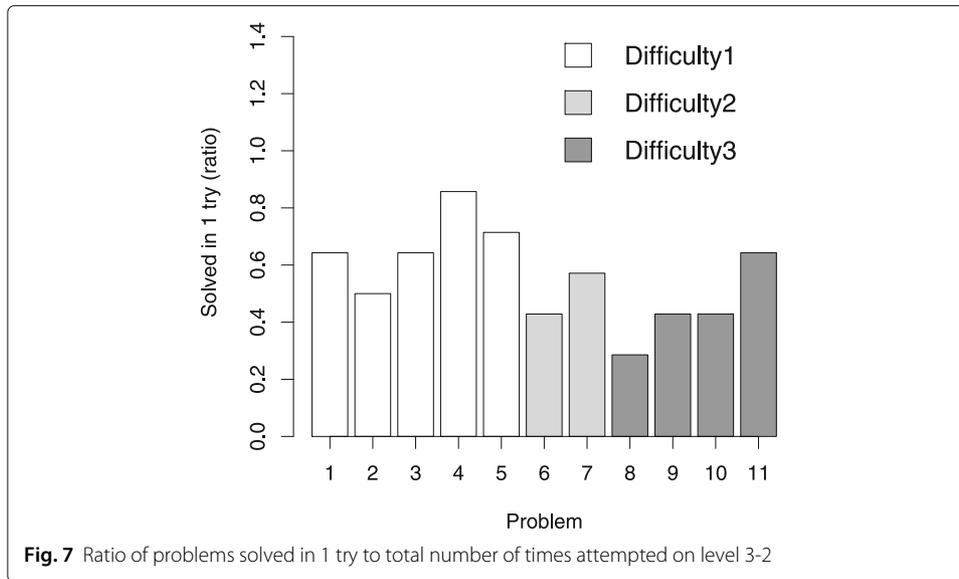


Fig. 7 Ratio of problems solved in 1 try to total number of times attempted on level 3-2

Case studies

The performance of four participants on level 3-2 is analyzed in this section. This means that their performance was close to trying to game the system. This could be a sign that they may not have engaged well with the system.

- 1 User 1: Performance of $1.0 \times$ GTS value. This participant had a lot of trouble in both peaks of difficulty (problems 6 and 8) and in problem 11. The participant quickly solved the other problems in difficulties 2 and 3.
- 2 User 2: Performance of $0.5 \times$ GTS value. This participant had trouble from problems 6 to 9. They quickly managed to solve problems 10 and 11.
- 3 User 3: Performance of $0.54 \times$ GTS value. This participant had similar results to user 1.

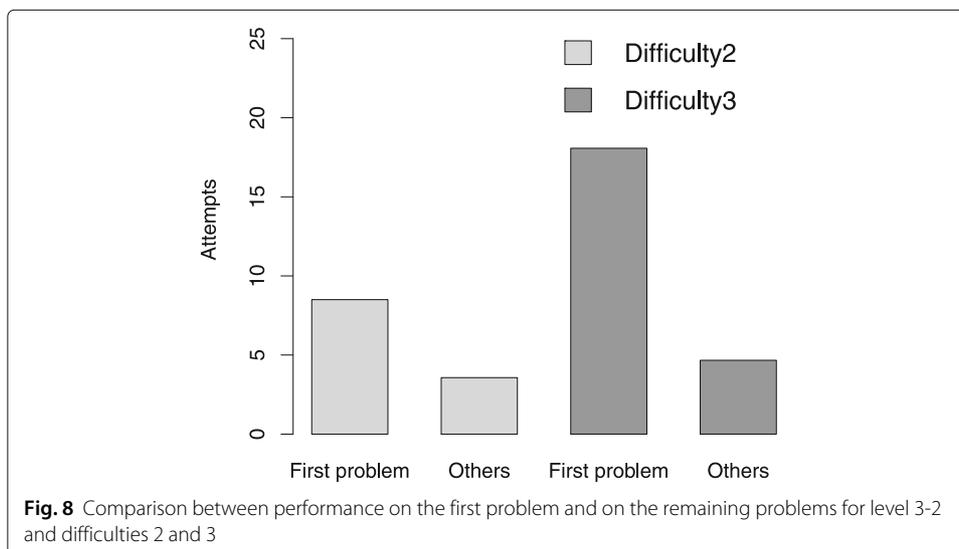


Fig. 8 Comparison between performance on the first problem and on the remaining problems for level 3-2 and difficulties 2 and 3

- 4 User 4: Performance of $0.84 \times$ GTS value. This participant had similar results to user 1.

All four students had good performance in previous levels. None of the students completed level 4. Since these students took time to get through level 3, the reason they did not complete level 4 is likely because of lack of time.

Three of the four users had trouble with problems 6, 8, and 11. They had more trouble than the other participants in the pilot study, and this pushed their average performance closer to GTS. That does not mean that they were not trying to actively solve and understand the problems since after going through this initial trouble, they managed to solve the other problems fast.

User 2 had a different scenario. The student had moderate trouble with problems 6 to 9, four problems in a row, and then the student's performance went up in problems 10 and 11. The student did not have as much trouble with problems 6 and 8 as the other participants, having more trouble with problems 7 and 9. Either the student finally understood the activity while trying to solve problem 9 or they received help from another student. We cannot be certain because although it was observed that some of the students were asking for help, it is not possible to know if the student corresponding to the examined log data asked for help or not.

Users had trouble with problem 11. Of the four problems of difficulty 3 in level 3-2, two of them are "decrease story" type, one is a "combination story" type, and the last one is an "increase story" type problem. Problems 8 and 11 are "decrease story" type problems, and participants had trouble with both. It could be the case that "decrease story" type problems are more difficult for students.

These results suggest that, even when students struggled at first, they could meaningfully take advantage of the application.

Results suggest that the interaction of students with the system was meaningful and that they showed growth patterns. This remains true even when we isolate the students that struggled at first. The overall analysis combined with the case studies answer our second research question, that the interactions with the Triplet Structure Model improve as the students use the application.

Limitations

The pilot study lacked a control group, a pre-test, and a post-test. This makes it hard to evaluate the learning effects of the application. Time constraints also limited the amount of data collected on the harder levels of the application.

Lastly, the interactions between students, and between teachers and students, may have influenced application use. In the future, one solution to this problem would be to isolate the children, so that data would be unaffected. Another approach would be to collect data on these interactions so that it could be compared to the collected software data. This would allow for a more complete view of application use and of the social implications of the application.

Conclusion

In this study, the design of an application for teaching the conceptual understanding of contextual problems for kindergarten students was presented. Furthermore, the viability

of the application was verified through the use of a kindergarten classroom. The interaction of users with the application was successful in many respects. Students showed good affective responses and satisfactory performance, outperforming the GTS calculated measures. The results suggest the following:

- 1 The application is successful in allowing for meaningful interaction with the Triplet Structure Model without the use of text and arithmetic expressions.
- 2 Students' interactions with the Triplet Structure Model improve while using the application, quickly clearing problems which had given them trouble in the past.

Their growth patterns and the distance to gaming the system patterns suggest that the interactions of the students with the software are meaningful. This indicates that the Triplet Structure Model is effective for younger children too. The results strengthen the software position as a tool to help kindergarten students to understand contextual problems conceptually.

The design and viability verification of the application, alongside the data analysis suggesting meaningful use and growth inside the application are the main contributions of this study.

The interactions of the students get better as they clear problems. As such, improvement inside the application has been verified. The lack of evaluations from outside the system, such as using a control group, pre-test, and post-test makes it hard to verify how well this transfer to other contexts. This is one issue to be tackled in future studies.

The case-by-case results suggest that students who struggled during some critical levels are also engaging and improving their understanding, they just took longer than other students to overcome hard challenges for the first time.

Looking at the average values, students perform well in the early levels. However, some of the students may be leaving level 1 and level 2 without a clear interpretation of the pictures. Adding user modeling and procedural question generation to the application is one approach to make sure that students are leaving the earlier levels with the desired interpretation of the pictures. Furthermore, increasing the number of easier problems in level 3 problems might make it easier for students to clear the harder levels.

The next step in this research is to allow students to create the story pieces themselves. The objective would be to make children think deeply about each quantity separately.

Abbreviations

CAI: Computer-assisted instruction; CSV: Comma separated values; GTS: Gaming the system; RPOT: Ratio of problems solved in one try

Acknowledgements

Not applicable.

Funding

This work was partially supported by JSPS KAKENHI Grant Numbers 17H01839 and 15H02931.

Availability of data and materials

Requests for data and materials will be evaluated on a case-by-case basis.

Authors' contributions

All authors contributed to designing and evaluating the application. PGFF developed the application, performed the statistical analysis, and wrote the manuscript. TH contributed in reviewing and discussing the manuscript. Both TH and YH contributed in integrating the design of the application with the Triplet Structure Model. KM contributed with his teacher expertise in improving the design of the problems and in improving the usability of the software for young children. KM also contributed in contacting the schools and in introducing the software to the children during the experiment. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹Graduate School of Engineering, Hiroshima University, 1-4-1 Kagamiyama, Higashi-Hiroshima 739-8527, Japan.

²Department of Management of Education, International Pacific University, Seto-cho Kanonji Temple, Okayama 709-0863, Japan.

Received: 15 June 2018 Accepted: 4 January 2019

Published online: 28 January 2019

References

- Baker, R, Walonoski, J, Heffernan, N, Roll, I, Corbett, A, Koedinger, K (2008). Why students engage in "gaming the system" behavior in interactive learning environments. *Journal of Interactive Learning Research*, 19(2), 185.
- Clements, DH, & Sarama, J (2007). Effects of a preschool mathematics curriculum: summative research on the building blocks project. *Journal for Research in Mathematics Education*, 38(2), 136–163.
- Dyson, NI, Jordan, NC, Glutting, J (2013). A number sense intervention for low-income kindergartners at risk for mathematics difficulties. *Journal of Learning Disabilities*, 46(2), 166–181.
- Freiman, V, Polotskaia, E, Savard, A (2017). Using a computer-based learning task to promote work on mathematical relationships in the context of word problems in early grades. *ZDM*, 49(6), 835–849.
- Furtado, PGF, Hirashima, T, Hayashi, Y, Maeda, K (2017). Learning arithmetic word problem structure with a picture combination application in kindergarten. *Proceedings of the 25th International Conference on Computers in Education, ICCE 2017*, 25.
- Gravemeijer, K, & Doorman, M (1999). Context problems in realistic mathematics education: a calculus course as an example. *Educational studies in mathematics*, 39(1-3), 111–129.
- Greer, B (1997). Modelling reality in mathematics classrooms: the case of word problems. *Learning and instruction*, 7(4), 293–307.
- Griffin, S (2004). Building number sense with number worlds: a mathematics program for young children. *Early childhood research quarterly*, 19(1), 173–180.
- Hamadneh, BM, Hamad, HA, Al-azzam, MM (2016). The impact of applying computer-based training strategy upon developing the skill of solving mathematical word problem among students with learning disabilities. *International Research in Education*, 4(1), 149–158.
- Hasanah, N, Hayashi, Y, Hirashima, T (2017). An analysis of learner outputs in problem posing as sentence-integration in arithmetic word problems. *Research and Practice in Technology Enhanced Learning*, 12(1), 9.
- Hegarty, M, Mayer, RE, Monk, CA (1995). Comprehension of arithmetic word problems: a comparison of successful and unsuccessful problem solvers. *Journal of educational psychology*, 87(1), 18.
- Hirashima, T, & Kurayama, M (2011). Learning by problem-posing for reverse-thinking problems, In *Artificial Intelligence in Education* (pp. 123–130). Berlin: Springer.
- Hirashima, T, Yamamoto, S, Hayashi, Y (2014). Triplet structure model of arithmetical word problems for learning by problem-posing, In *International Conference on Human Interface and the Management of Information* (pp. 42–50). Berlin: Springer.
- Hirashima, T, Yokoyama, T, Okamoto, M, Takeuchi, A (2007). Learning by problem-posing as sentence-integration and experimental use, In *AIED*, vol. 2007 (pp. 254–261). Berlin: Springer.
- Huang, T-H, Liu, Y-C, Chang, H-C (2012). Learning achievement in solving word-based mathematical questions through a computer-assisted learning system. *Educational Technology & Society*, 15(1), 248–259.
- Jordan, NC (2010). Early predictors of mathematics achievement and mathematics learning difficulties, In *Encyclopedia on Early Childhood Development* (pp. 11–15). Retrieved from <http://www.child-encyclopedia.com/sites/default/files/dossiers-complets/en/numeracy.pdf>
- Koedinger, KR, & Anderson, JR (1998). Illustrating principled design: the early evolution of a cognitive tutor for algebra symbolization. *Interactive Learning Environments*, 5(1), 161–179.
- Larman, C, & Basili, VR (2003). Iterative and incremental developments. A brief history. *Computer*, 36(6), 47–56.
- Rivera, F (2014). *Teaching to the math common core state standards: focus on kindergarten to grade 5*. Berlin: Springer.
- Starkey, P, Klein, A, Wakeley, A (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*, 19(1), 99–120.
- Supianto, AA, Hayashi, Y, Hirashima, T (2017). Model-based analysis of thinking in problem posing as sentence integration focused on violation of the constraints. *Research and Practice in Technology Enhanced Learning*, 12(1), 12.
- Wilson, AJ, Dehaene, S, Dubois, O, Fayol, M (2009). Effects of an adaptive game intervention on accessing number sense in low-socioeconomic-status kindergarten children. *Mind, Brain, and Education*, 3(4), 224–234.
- Yamamoto, S, Kanbe, T, Yoshida, Y, Maeda, K, Hirashima, T (2012). A case study of learning by problem-posing in introductory phase of arithmetic word problems, In *Proceedings of the 20th International Conference on Computers in Education* (pp. 25–32). Singapore: ICCE 2012 Organizing Committee.
- Yamamoto, S, Akao, Y, Murotsu, M, Kanbe, T, Yoshida, Y, Maeda, K, Hayashi, Y, Hirashima, T (2014). Interactive environment for learning by problem-posing of arithmetic word problems solved by one-step multiplication and division, In *Proceedings of the 22th International Conference on Computers in Education* (pp. 89–94). Japan: ICCE 2014 Organizing Committee.
- Yerushalmy, M (2006). Slower algebra students meet faster tools: solving algebra word problems with graphing software. *Journal for Research in Mathematics Education*, 37(5), 356–387.