Enhancing young children's reasoning about data distribution through model creating and sharing

Takashi Kawakami, Nishikyushu University, Japan

1. Introduction

Distribution is at the heart of statistics and is a fundamental component of statistical reasoning, hence it is very important to develop students' notions of distribution in future school statistics curricula (Burill and Biehler, 2011). The necessity of fostering students' informal views of distributions from earlier ages arises because of the difficulties of developing an aggregate view of distribution (e.g., Ben-Zvi and Amir, 2005). Recent research has indicated that modeling activities related to data provide an effective means for illuminating and promoting young children's statistical reasoning (e.g., English, 2012, 2013). How can activities incorporating the idea of models and modeling develop young children's informal views of distributions? The aims of this paper are (a) to illuminate the development of Year 1 children's informal views of distributions through model creating and sharing and (b) to derive some suggestions for quality teaching and learning of statistics in the early years of schooling.

2. Developing children's views of distributions by utilizing the idea of models and modeling

Distribution is a complex and multifaceted entity. Bakker (2004) explains that distribution has various aspects such as centre, spread, density, skewness, and so on, from the perspective of didactical phenomenology. Based on the ideas of Bakker, this study intends to enable students to reason statistically with an aggregate view of data by connecting the elements of distribution, such as recognising patterns in a data set.

Several researchers have indicated that young children have the potential to develop informal ideas related to distributions (e.g., Ben-Zvi and Amir, 2005). Modelling that includes creating and revising models is increasingly highlighted as a powerful vehicle for developing young children's mathematical ideas and higher-order competencies (Lesh et al., 2013). Likewise, modelling can provide young children with access to fundamental statistical ideas such as data, variability and distribution (e.g., English, 2012, 2013; Lehrer and Schauble, 2004).

A *model* is defined as a representation of structure in a given system from the cognitive perspective (Hestenes, 2010). In addition, Hestenes (2010) defines the *system* as a set of related objects, and explains that the *structure* of a system is a set of relations among the objects in the system. Using this notion of a model, it is crucial that students represent their internal models (such as images, where the structure of a situation is reflected), using notation to create external models that can be shared with others (Kawakami and Matsuzaki, 2012). In this study, two activities involving creating models and sharing them are set up in Year 1 lessons of pictographs. The first involves children creating and revising distribution-related pictographic models from a data set. The second involves sharing and discussing classmates' distribution-related models. Although several researchers have examined how Year 1 children made pictographs (English, 2012; Matsumoto, 2010; Watson and Moritz, 2001), little research exists on

developing children's statistical reasoning about distributions when making pictographs, by incorporating the idea of models and modelling.

3. Methodology

The participants

The study was conducted in a private primary school in Tokyo by the author. The experimental class consisted of two lessons (45min×2) and contained 27 Year 1 children (19 male, 8 female) aged 6-7 years old. The children had not had statistical experiences at school or learned about pictographs before.

Research design

This study involves one instance of participatory lesson study of planning, action and observation, and reflection designed to improve professional practice in the teaching and learning of primary mathematics. After the first lesson, a post-lesson discussion was implemented, in which the teacher and participants shared data from the lesson to better understand children's learning, and reflected on the task, teacher's questioning, and activities in preparation for the next lesson.

Task and activities

Library Task

This school library map (Figure 1) shows the books of each five older students (Year 1-Year 5 students at that time) borrowed last January. How can you move the book cards and the student cards on the map on your board, so that your friends can quickly find the difference between the number of books each older student borrowed?

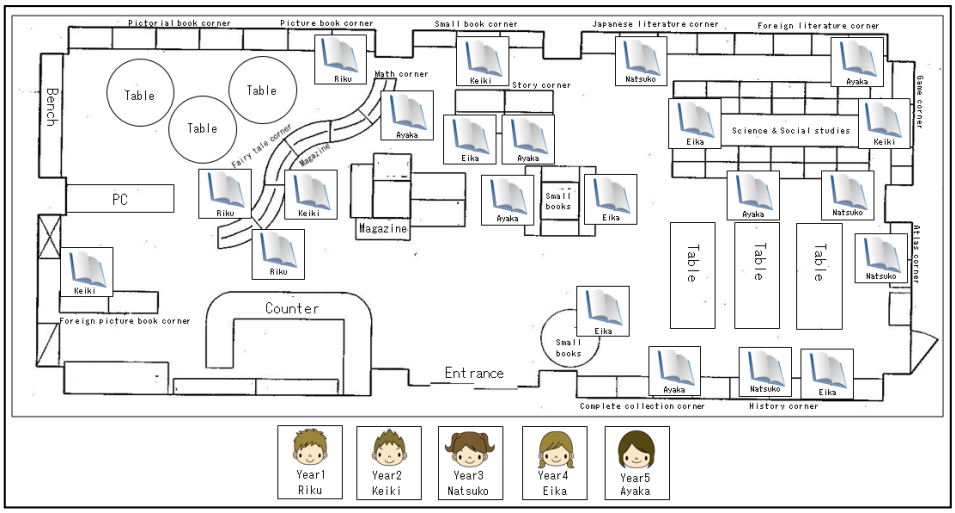


Figure 1. School library map with student cards and book cards (translated by English)

The context of the *Library Task* used in the lessons was "encouragement of reading." Reading is a part of school life. The context was chosen so that young children can have access to data about older students' reading. The need to display data was emphasized from the viewpoint of model sharing. Furthermore, each of the children was supplied

with one mini white board and one marker, and mini magnet cards depicting 22 books and five older students were set on the school library map (Figure 1), so that the children could externalize and reflect their distribution-related models, as shown in Figure 2.

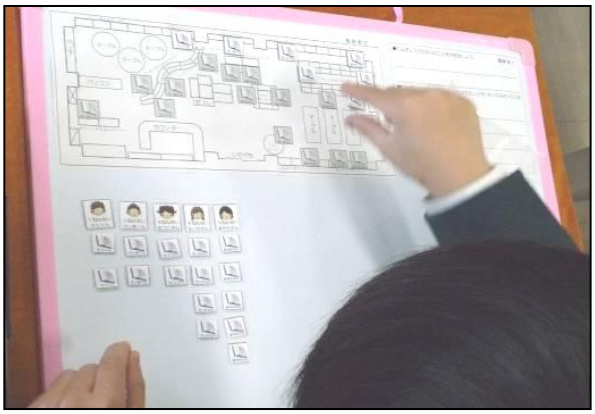


Figure 2. White board with school library map and magnet cards

In the first lesson, the children represented the distribution of the number of borrowed books. At first, the teacher asked the children whether older students borrowed more books than them last month, and showed the same map as Figure 1. The map was a simplified and symbolized the model to some extent, as the teacher intended to focus on the activity of creating graphical representations. The teacher demonstrated the correspondence between the map and photos of the library, and explained the map concretely (e.g., "The map shows that Riku borrowed a book in the picture book corner"). Next, the children were encouraged to attempt the *Library Task* by moving the cards on the board.

From the end of the first lesson to the beginning of the second lesson, they shared their models of the data, and discussed better representations for showing the difference in the number of borrowed books. Furthermore, they were asked to add their own values of the number of borrowed books, and improved their initial models and compared the older students' data to their individual values. At the end of the second lesson, they set a target about reading by referring to their improved models.

Data collection and analysis

The data collection comprised audio and video recordings, lesson artefacts and field notes. Protocols were made from the audio and video data. The artefacts at the focus of analysis included digital images of children's graphical representations and comments through each lesson as a supplementary explanation of the models. In the analysis, Watson and Moritz (2001)'s levels for representing data were adopted. The levels are identified by the fact that learners are able to take account of and operate, which is based on the cognitive development framework of Biggs and Collis (1982). This study analysed what elements of distribution (e.g., centre, spread, density, shape) and/or context of the distribution the children focused on and how they connected the elements

through two lessons, in order to illustrate how the children's original models were maintained or changed/modified by sharing them with their classmates.

4. Selected findings

Children's Initial Reasoning in Lesson 1

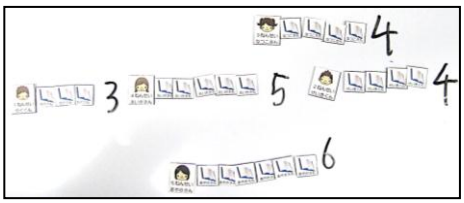
Table 1 shows how 27 children reasoned statistically when attempting the *Library Task* in the first lesson, together with typical examples of children's models. Three types of models were distinguished from the viewpoint of included elements of distributions.

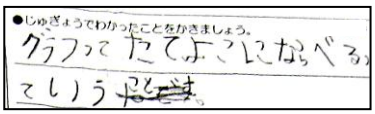
Firstly, only one child grouped the book cards for each student card in a row without using a baseline (*Model A*). This representation included only a single aspect of the distribution, the frequency of books (*density*), as shown in Fusa's model. Fusa perceived only a part of the strength of graphs, as shown in his comment. Secondly, 23 children used a baseline of the student cards arranged in a single row or column, using this as a baseline, with the book cards for each person arranged in a single column or row, respectively (*Model B*). In most of *Model B*, the student cards were ordered according to the grade and/or frequency, but patterns were not described. These representations included several aspects of the distribution, but did not connect the aspects. For example, it may be gathered from Taku's comment that his model showed only the frequency of books (*density*) and the difference of the frequency (*density*) with one to one correspondence of the book cards. Lastly, three children also paid attention to patterns in the distribution (*Model C*). Their representations included the connection between aspects of the distribution, such as the functional relation between the category of students' grades (*spread*) and the frequency of books (*density*), as indicated in Shun's comment.

Table 1. Children's models in lesson 1 (N=27)

Types of Models & Examples	Frequency
Model A: Representation of a single aspect of the distribution, such as grouping	1

a. Fusa's model





[Findings from the lesson]
A graph is arranged in length or horizontally.