HOW DO PRIMARY SCHOOL STUDENTS BEGIN TO REASON ABOUT DISTRIBUTIONS?

DANI BEN-ZVI
University of Haifa
dbenzvi@univ.haifa.ac.il

YAEL SHARETT-AMIR
University of Haifa
amir@maaganm.co.il

ABSTRACT

This study explores the emergence of second graders' informal reasoning about distribution in a carefully planned learning environment that includes extended encounters with open-ended Exploratory Data Analysis (EDA) activities. The current case study is offered as a contribution to understanding the process of constructing meanings, language, representations and appreciation for distributions at an early age of schooling. It concentrates on the detailed qualitative analysis of the ways by which three second grade students (age 7) started to develop views (and tools to support them) of distributions in investigating real data, inventing and using various informal data ideas and representations. In the light of the analysis, a description of what it may mean to begin reasoning about distribution by young students is proposed, and implications to teaching, curriculum and research are drawn.

Keywords: EDA; Data; Distribution; Statistical reasoning; Conjecture; Growing a sample

1. OVERVIEW OF THE PROBLEM AND ITS IMPORTANCE

Having observed statistics lessons in primary classrooms in Israel, we became concerned by the fact that statistics is frequently portrayed in a very narrow and limited way, which can be encapsulated: "every phenomenon can be captured by a bar chart". Starting in kindergarten, early statistics instruction focus mostly on skills, describing (sometimes real) phenomena in conventional graphs (how to draw, how to interpret), and later (fourth or fifth grade) on calculating averages. This limited view of statistics as a mathematical discipline is imposed partly by teachers' limited statistical knowledge as well as by the curricular materials they use. We have had a sense that if we only allow young students construct and express their curious and creative ideas about data, we would be surprised to learn how much they are capable of doing and understanding.

These situation and expectations motivated us to try providing young students with supportive, carefully planned, but open, learning environment that aims at gradually building students' confidence to experiment with data and complex statistical notions such as distribution, samples, and variability. In this study we attempt to "push" the notion of distribution to early age, i.e., second grade students, in order to look for any (even tiny) indication for students' potential to develop their statistical skills and reasoning about data and distribution.

Distribution is at the heart of statistics and is a fundamental component of statistical thinking (Bakker & Gravemeijer, 2004). In his key address to the SRTL-4 Research Forum, Wild (2005) – past president of the IASE – proposed that the statistical response to the "omnipresence of variability" (Cobb & Moore, 1997) is to investigate, disentangle and model patterns of variation in order to learn from them. Statisticians look at variation through a lens which is "distribution". A distribution is "the pattern of variation in a variable" (or set of variables). Having collected a sample of data and set aside their labels, what we are left with is displays or summaries of distributions (Wild, 2005). Distributions – the basic core materials of exploratory data analysis (EDA) work – are used by statisticians and learners as "clay in the hands of the potter".

Distributions come in different meanings, names, contexts, levels of abstraction, shapes, and uses (e.g., a distribution of a variable, normal distribution, conditional and marginal distributions, binomial distribution, sampling distribution, and the chi-square, *t* and the *f* families of distributions). The three main contexts in which statisticians use the term are: experimental distribution, theoretical distribution, and sampling distributions (Wild, 2005). The multi-faceted distribution should therefore be repeatedly integrated, revisited and highlighted in statistics curriculum and instruction. Research however provides us with a strong case that understanding of distributions is much more complex and difficult than prior literature suggests. For example, students or teachers, examined in some of the studies, demonstrate diverse intuitions, misconceptions, and incomplete or shallow understanding of distributions (Garfield and Ben-Zvi, in review).

Experiments to support students reasoning about distribution were focused mainly in the middle school level, and partly in the last years of primary school (grades 5-6). Cobb (1999) and McClain & Cobb (2001) focused on the overarching concept of distribution in their instructional and research efforts in seventh grade. So did others at the middle school level (e.g., Ben-Zvi & Arcavi, 2001; Bakker, 2004). *Tinkerplots*, an innovative technological tool, provides unusual pedagogical affordances for upper primary students to experiment with data and distribution, using their basic skills to sort, stack and separate data. However, our claim is that distribution – as a basic building block in statistical reasoning – should be centrally emphasized from earlier ages (grades 1-3) in formal and informal activities and discussions. Suggestions for how this might be encouraged and evaluated at the primary level are rare in the current literature.

How early can we start pondering ideas related to distributions? This question can be rephrased in several ways: What skills, knowledge and intuitions do kids at an early age bring with them to data explorations that may support understandings (even partial) of distribution? What language, discourse, considerations and arguments are used by them when they deal with data, data representations, and distributions in a challenging problem-based environment? What is the role of context (the real world situation, the background knowledge) in supporting their construction of understanding of the distribution at hand? What are good starting points, pedagogical "tricks' such as concrete artifacts and appropriate tasks that work with youngsters? These are all puzzling and important questions if we believe data and chance are to be introduced at an early age.

Why are these questions important? Imagine we could understand the basic foundations of students' naïve, intuitive, and informal statistical reasoning and discourse. We then would inform and focus our curricular and pedagogical efforts when data and chance are introduced to students at early age (as highly recommended by the current statistics education reform movement, e.g., DfES, 2000; NCTM, 2000). If we knew what students already know, and what they almost know (in the sense of the Zone of Proximal Development, Vygotsky, 1978), our learning trajectories (Cobb & McClain, 2004) can create an arena for fuller expressions of students' curiosity, creativity and abilities, while they invent and acquire simple and effective tools of statistical investigations. Instead of telling young students what distributions are and what type of data representations are correct to use, we would encourage them first to organize and discuss their data sets and represent them in innovative ways, learn by doing what works for them and what doesn't and why, what works better and why. Thus, this study is suggested as a contribution towards supporting and guiding a bottom up approach to instruction and learning data and chance (Konold, 2002).

The current study explores the emergence of young students' informal reasoning about distribution in a carefully planned learning environment that includes extended encounters with open-ended EDA activities. We share with many others in the statistics education community the belief in the ideas of EDA as an important starting point for students. The current extended case study is offered as a contribution to understanding the process of constructing meanings, language, representations and appreciation for distributions.

We start with a brief literature review focusing on reasoning about distribution and the value of comparing data sets tasks, followed by a description of setting and context. We then concentrate on the detailed qualitative analysis of the ways by which three second grade students (age 7) started to develop views (and tools to support them) of distributions in investigating real data, inventing and using various informal data ideas and representations. In the light of the analysis, a description of what it may mean to begin reasoning about distribution by young students is proposed, and implications to teaching, curriculum and research are drawn.

2. BRIEF REVIEW OF RELATED LITERATURE

On the one hand, distribution has a complex and multi-faceted structure that requires comprehensive knowledge in mathematics and probability as well as abstract reasoning. On the other hand, distribution is a basic part of a larger conceptual structure of statistics consisting of a web of big ideas such as variation and sampling (Reading & Shaughnessy, 2004; Watson, 2004). Therefore it has been suggested to deal informally and coherently with all these big ideas at the same time with distribution in a central position. As Cobb (1999) proposes, focusing on distribution as a multifaceted end goal of instruction might bring more coherence in the statistics curriculum. There is some evidence however that students at various ages tend to conceive a data set as a collection of individual values instead of an aggregate that has certain properties (Bakker & Gravemeijer, 2004; Ben-Zvi & Arcavi, 2001; Ben-Zvi, 2004; Hancock, Kaput, & Goldsmith, 1992; Konold & Higgins, 2003). Despite these observations, statistics educators and researchers recommend guiding students how to reason about distributions in several informal ways, e.g., through the investigation of frequencies distributions, focus on distributional shape and modal clumps (e.g., Bakker & Gravemeijer, 2004; Cobb, 1999; Cobb & McClain, 2004; Petrosino, Lehrer & Shauble, 2003).

Bakker & Gravemeijer (2004) suggest several promising instructional heuristics too support students' aggregate reasoning of distributions: (1) Letting students invent their own data sets could stimulate them to think of a data set as a whole instead of individual data points. (2) *Growing samples*, i.e., letting students reason with stable features of variable processes, and compare their conjectured graphs with those generated from real graphs of data. (3) Predictions about the *shape* and location of distributions in hypothetical situations. All these methods can help students to look at global features of distributions, foster a more global view, and see *the signal in the noise* (Konold & Pollatsek, 2002).

Bakker and Gravemeijer (2004) also offer an analysis of the relation between data and distribution. They suggest a three level structure: data as individual values, aspects of both data sets and distributions (such as center and spread), and distribution as a conceptual entity. Students tend to see individual values, which they can use to calculate, for instance, the mean, median, range, or quartiles. This does not automatically imply that they see mean or median as a measure of center or as representative of a group (Mokros & Russell, 1995; Konold & Pollatsek, 2002). In fact, students need a notion of distribution before they can sensibly choose between such measures of center (Zawojewski & Shaughnessy, 2000). Therefore, students need to conceive center and spread as characteristics of a distribution, and looking at data with a notion of distribution as an organizing structure or a conceptual entity. Bakker and Gravemeijer (2004) suggest that reasoning with shapes forms the basis for reasoning about distributions. In this study we question young students' basis for reasoning about distributions by using a variety of methods including comparing groups activities.

Comparing groups provides the motivation and context for students to consider data as a distribution and take into account and integrate measures of variation and center (Konold & Higgins, 2003). There is some evidence however that the group comparison problem is one that students do not initially know how to approach and the challenge may remain even after extended periods of instruction. Students' difficulties may stem from the multifaceted knowledge necessary for comparing groups, such as understanding distributions (Bakker & Gravemeijer, 2004), understanding averages (Bright & Friel, 1998; Gal, Rothschild & Wagner, 1990; Hancock, Kaput & Goldsmith, 1992; Konold, Pollatsek, Well, & Gagnon, 1997; Watson & Moritz, 1999), representativeness (Mokros and Russell, 1995), and variability in data (e.g., Meletiou, 2002). Students also have difficulties in adopting statistical dispositions, such as tolerance towards variation in data, and integration of local and global views of data and data representations (Ben-Zvi & Arcavi, 2001; Ben-Zvi, 2002; Ben-Zvi, 2004). Cobb (1999) proposes that the idea of middle clumps (“hills”) can be appropriated by students for the purpose of comparing groups. Almost none of these studies have dealt directly with very young students, which are the focus of the current study.

3. PURPOSE/GOALS OF THE STUDY AND ITS SPECIFIC RESEARCH QUESTIONS

These studies on older age students, lack of studies on the younger age, and suggested statistics learning goals formed the motivation to explore the possibilities for students in early primary education (grade 2, age 7) – with little or no prior statistical knowledge – to develop an informal understanding of distribution in a variety of EDA situations. The following research question is used to structure the current study and the analysis of data collected: *How do second grade students begin to reason about distribution in a rich and supportive EDA learning environment?* Such a learning environment involves peer collaboration and group discussions, open-ended EDA tasks and minimal guidance of a teacher/researcher. In the current paper, we focus on particular type of tasks: conjecturing and "growing a sample" in describing one data set or comparing distributions situations. The goal is to closely trace the emergence of beginners’ reasoning about distribution:

- How do young students come to reason about distributions?
- What are their intuitive emerging conceptions of distributions?
- What is the role of representations they use in the emergence of their reasoning about distribution?
- What is the role of the conjecturing and "growing a sample" instructional activities?

4. CONTEXT AND METHOD

The current study follows closely the behavior and discourse of three second grade students engaged with relatively simple EDA tasks. It took place in a Kibbutz[1] primary school in Israel. Skillful and experienced teacher (one of the researchers, a graduate student in mathematics education) observed and gently guided the students during ten informal sessions (about one hour each). In these sessions, students experienced some of the processes involved in genuine data-based enquiry. The students were asked to study and investigate data related to losing deciduous teeth in five grade levels in their school (kindergarten to fourth grade). In particular interest of the current report, they were asked to conjecture about expected values before they collected sample data, and to engage in "growing a sample" after they collected data. They were encouraged to discuss their ideas, invent and use multiple ways of representing the distributions and to reflect on them. It was expected that these data investigations and the whole learning environment and trajectory would support and partially reveal the development of beginners' informal reasoning about distribution.

4.1 APPROACH

The primary goal of this study is not to assess the effectiveness of the instructional design, but rather to study the emerging statistical reasoning of the students (the experiment population) along a planned learning trajectory provided by the researchers. To address the main questions, we carried out design research (also called developmental research, see Freudenthal, 1991; Gravemeijer, 1994; Edelson, 2002; Cobb & McClain, 2004; Ben-Zvi, Garfield & Zieffler, in press). The research had three stages that will be elaborated below: the preparation phase, the actual experimentation phase, and the retrospective analysis (Gravemeijer, 2000; Steffe & Thompson, 2000).

4.2 "PARTICIPANTS" OR "SUBJECTS"

Three second grade students (age 7) took part in the study: Yonat and Irit (females) and Omri (male). They were selected by their educator who indicated that they were generally average students but with high ability to concentrate, persist and communicate. The students engaged seriously and enthusiastically with the activities offered by the researcher, trying to understand and offer their solutions and ideas, and were quite independent in their work. The fact that they were videotaped did not intimidate them.

---

[1] A rural community in Israel based on communal property, in which members have no private property but share the work and the profits of some collective enterprise, typically agricultural but sometimes also industrial.

One of the main issues that have to be addressed when relatively complex concepts are introduced to young children is their limited abstraction capabilities. Young children have a difficult time verbalizing their thoughts, especially when it concerns abstract concepts and actions (Piaget 1971, 1973). While children can be extremely honest in their feedback and comments concerning their environment and actions, much of what they say needs to be interpreted within the context of concrete experiences (Druin, 1999). We were aware of these difficulties in all stages of the study, and provided many types of scaffolding to serve two purposes: to maximize students functioning in the suggested learning environment and to assist in revealing their reasoning processes. Scaffolding took the forms of using extremely familiar and relevant context, whole numbers data ranging 0-20, use of concrete instructional materials (e.g., cards), encouragement of group interactions and discourse, repeated cycles of investigations, and ongoing reflection on actions and artifacts.

4.3 TASKS

Context

The phenomenon of losing milk teeth provides rich context for this study. While the teeth of most vertebrates are replaced continuously throughout the animal's life, the pattern found instead in most mammals is *diphyodonty*, a term derived from Greek words meaning "twofold production of teeth." Humans have two sets of teeth. The teeth that appear first are called the *milk teeth* or *deciduous dentition* or the primary teeth. These teeth are later shed off and are replaced by permanent set of teeth called the *permanent dentition*. Milk teeth are called so due to their white color which resembles the color of milk. The milk teeth are whiter than the permanent teeth which replace them.

Milk teeth play an important role in the alignment and spacing of permanent teeth. They support the upper and lower jaws vis-à-vis each other. Constant chewing with deciduous dentition is an essential prerequisite for the correct development of the jaws. If milk teeth are lost too early on, the position of the other teeth may be affected, which is why it is important that milk teeth should remain intact until they are replaced by permanent teeth.

The 20 milk teeth are shed from about 5-6 years onwards till about 10-12 years of age. These teeth are pushed out by permanent teeth growing behind the milk teeth. The permanent teeth start erupting by about 6 years of age and most of them erupt by about 12 years of age except the third molars or the wisdom teeth, which erupt any time between 18-25 years of age. Thus, between the ages of about 6-9 years the child has some milk teeth as well as some permanent teeth. This period is called the mixed dentition period.

Instructional activities

In line with the principles of the National Council of Teachers of Mathematics (NCTM, 2000), we looked for ways to guide students in being active learners dealing with increasingly sophisticated means of support. Activities were designed to be based on students' own ideas and not to guide them toward more conventional notions and representations. Students were involved in planning for data collection, sample data collection, representing the collected data (inventing graphs or organizing prepared artifacts to create graphs), interpreting real and hypothetical data, making *conjectures* about possible outcomes, and "*Growing a sample*" activities. Students' responses to the last two types of activities in a univariate data set and comparing distributions tasks are the focus of the current study.

In a conjecture task, students were asked to suggest few hypothetical values of the variable of a specific age group *before* they actually collected data from a sample of that age group. In "growing a sample" task students were asked to predict additional values of the variable *after* they had collected sample data. Similarily, Konold & Pollatsek (2002) and Bakker and Gravemeijer (2004) asked students to predict and explain what happens to a graph when bigger samples are taken. They claim that this type of activities have the potential to foster students understanding of distributions as a global feature by making them concentrate on the patterns in the variable processes.

A variety of pedagogical and methodological techniques were used during the activities, such as, students reporting to other second grade students who were not part of this study and to their parents, discussing and reflecting on students' created artifacts, individual interviews with the researcher, and a combination of individual and group tasks. The following table (Table 1) presents the main activities during the ten sessions.

Few weeks before this study, we conducted a pilot study to test whether students at that age are able to work with the milk teeth data in multiple and meaningful ways. Results were positive but also indicated difficulties to discuss and explain their ideas and actions. Therefore, instructional activities were redesigned in the current study to focus more on "doing".

4.4 ANALYSIS

The research team performed a retrospective analysis after each session (to re-direct next session learning trajectory), as well as after the entire teaching experiment has been completed. Students' responses, behavior and gestures were analyzed using video recordings of all the sessions and their full transcripts, researcher's observations, students’ interviews and artifacts. The analysis of the videotapes is based on interpretive microanalysis (see, for example, Meira, 1991): a qualitative detailed analysis of the protocols, taking into account verbal, gestural and symbolic actions within the situations in which they occurred. The goal of such an analysis is to infer and trace the emergence and development of cognitive structures and the socio-cultural processes of understanding and learning.

The interpretive microanalysis is one type of microgenetic methods which are essential for answering questions about how learning occurs. Microgenetic methods have three main properties: (1) Observations span the period of rapidly changing competence; (2) within this period, the density of observations is high, relative to the rate of change; and (3) observations are analyzed intensively, with the goal of inferring the representations and processes that gave rise to them (Siegler and Crowley, 1991).

Two stages are used to validate the analysis, one within the researchers’ team and one with additional researchers in mathematics education, who have no involvement with the data (triangulation in the sense of Schoenfeld, 1994). In both stages the researchers discuss, present, and advance and/or reject hypotheses, interpretations, and inferences about the students’ cognitive structures. Advancing or rejecting an interpretation requires: (a) providing as many pieces of evidence as possible (including past and/or future episodes, and all sources of data as described earlier) and (b) attempting to produce equally strong alternative interpretations based on the available evidence. The report includes cases in which the two analyses are not in full agreement, and points of doubt or rejection are not refuted or resolved by iterative analysis of the data.

Table 1: Topics and main activities of sessions.

#	Session Topic	Student's Activities
1.	Introducing the investigation and first data collection	• Introductory discussion about deciduous teeth. • Planning data collection from second grade students. • Conjecturing on kindergarten students and third graders. • Collecting data from a sample of third graders using forms they created. • Conjecturing about kindergarten students and second graders.
2.	Representing data and growing a sample	• Representing third grade data (individual task). • Growing the sample of third grade.
3.	Collecting more data and conjecturing	• Reflecting on student's representations (personal interview). • Conjecturing on second graders and drawing sketches of hypothetical distributions. • Preparing data collection forms for second grade. • Collecting second grade data.
4.	Representing data and conjecturing	• Comparing second grade sample with hypothesized data. • Conjecturing on K-4 and drawing sketches of hypothetical distributions. • Representing second grade data (individual task).
5.	Revisiting issues	• Reporting to other students about their work. • Conjecturing on 10 kindergarten students and drawing sketches of hypothetical distributions. • Preparing forms for data collection from first grade classes.
6.	Collecting data	• Collecting first grade data.
7.	Comparing distributions 1 (group)	• Comparing graphically between first and second grade distributions using two-color cards (group task).
8.	Comparing distributions 2 (individual)	• Interpreting the previous session comparing groups' representations. • Comparing graphically between first and second grade distributions using two-color cards (individual task).
9.	Comparing distributions 3 (group)	• Comparing graphically between first and second grade distributions using gradient-color cards (group task).
10.	Reflection and summary	• Discussing what they have learnt with their parents.

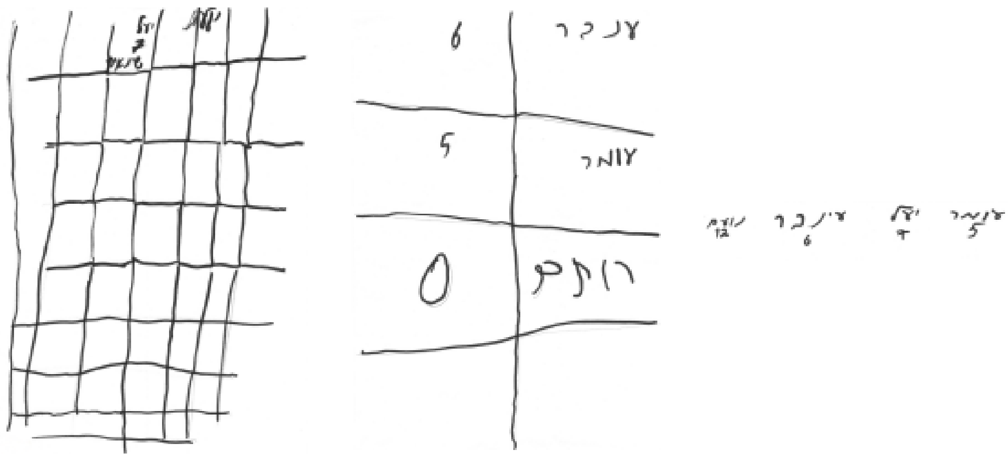
5. RESULTS

During the first five sessions of the study, the researcher suggested six tasks involved with making a conjecture or "growing a sample" to reveal if the students understand the context of the problem; if and how they bring to an expression the notions of frequency, distribution and variability in the data; how they handle zero or missing values; and how they organize and represent their data. The description will include the following sections: First data organizations ("tables"); first conjecture: How many teeth do kindergartners and third graders lose? First "growing a sample" activity: what if you meet additional six third graders? and comparing hypothetical distributions.

5.1 FIRST DATA ORGANIZATIONS: "TABLES"

The first session started with a short introduction and discussion about the shedding of deciduous teeth in animals and humans, in which the students showed their great interest and motivation in studying the topic. The researcher took their initial responses as an indication that the context is familiar enough to start with the first EDA task. She asked the students to investigate the phenomena of falling teeth in second grades of their own school (two classes, 56 students). In response, the students spontaneously talked about the number of teeth they personally lost (Yonat – 7 teeth, Irit – 6, Omri – 5) and the pleasure of getting presents from their parents when they lose a tooth. When asked how they would know about the rest of their friends in second grade, Irit suggested right away drawing a "table". The researcher asked each one of them to draw his ideas of how to collect and organize the second grade data on a separate piece of paper (Figures 1a-c).

Yonat drew a grid and wrote her own data ("Yonat 7 teeth") in the center of the first row (Figure 1a). She also suggested writing the sum of all numbers in one of the cells at the bottom of the table. Irit organized the data in a more conventional structure: Two vertical columns for student's name and number of falling teeth. She also entered real data (of herself, Omri and her kindergarten sister - Rotem) at the top cells of her table (Figure 1b). Omri's organization was similar to Irit's but is laid horizontally and does not have separation lines (Figure 1c). He indicated that additional data would have to be added in rows below the first ones (creating a grid). While the first inscription (Figure 1a) included a label ("teeth") near the value seven, the other two presented just the numbers without labels, which were considered unnecessary by the students.



a. Yonat *b. Irit* *c. Omri*

Figure 1: First "tabular" data organizations

In this task, the students spontaneously suggested their intuitive ideas of tabular forms to collect data. In theses inscriptions, data is stacked and separated in a fairly organized manner. We distinct between two table types: (a) *Cards table* (Figure 1a) – resembles organization of data on cards structured one by the other in a crisscross grid; and (b) *Semi-conventional table* (Figures 1b and 1c) – separate among variables, and avoid the variable label in each cell. These two basic structures were also observed in other first and second grade EDA classes, and were typical in this study to students' preliminary data inscriptions used for collecting and recording data.

Yonat's explicit suggestion to include the *total* in her inscription (Figure 1a) is typical to students' initial inclination in this study to view the inquiry's purpose as arithmetic: finding the teeth total of second graders. A distribution in this view is not about describing a general shape of a data set, but rather summing up the numbers. Considering the total as an important purpose had a conflicting role on students' reasoning: It helped develop students' statistical skills, such as cleaning data by deleting duplicate values, but also hindered the development of aggregate view of distribution.

5.2 FIRST CONJECTURE: HOW MANY TEETH DO KINDERGARTNERS AND THIRD GRADERS LOSE?

Before they collected their first set of data from a sample of kindergartners[2], the following discussion took place[3]:

351	*Researcher*	Each one of you will ask about six kindergarten kids …
352, 355	*Yonat*	But teeth do not fall in the kindergarten! … My sister is almost the oldest in the kindergarten and did not lose even one tooth!
356	*Researcher*	Yes?
357	*Irit*	I have lost two [teeth] in the kindergarten.
358	*Researcher*	See, Irit lost two.
359	*Yonat*	I have lost one in the kindergarten.
360	*Omri*	I also lost two in the kindergarten.
361	*Researcher*	Which numbers then do you roughly think they [the kindergartners] will tell you?
362	*Yonat*	Two and one.
363	*Irit*	[pointing out fingers] One, two, three.
364	*Yonat*	Not three!
365	*Omri*	Not much.
366	*Yonat*	Omri has almost lost three [teeth].
367	*Researcher*	Ahah …
368	*Irit*	But, we have lost almost three teeth [when we were kindergartners].
369	*Yonat*	I [still think that the maximum is] two.

(*First session, 351-369*.)
(To view this discussion, see Video Segment 1*: First conjecture: How many teeth do kindergartners and third graders lose?*)

This dialogue about the expected values in the kindergarten is initiated by Yonat's surprising observation: "*But in the kindergarten teeth do not fall*" [*352*]. She reinforced her argument by providing an example of her sister, who happened to be among the oldest kids in the kindergarten group that did not lose any teeth. However, the other students disagreed and provided counterexamples – their own cases (Irit and Omri – two, and Yonat – one). When directly asked by the researcher about the expected values, the students agreed that one and two teeth are probable in the kindergarten sample, but disagreed on the suggested maximum value – three teeth. They used a mix of arguments relating to age-teeth correlation and their personal experiences. Yonat used Omri's case to support her claim about the supposed absence of three teeth in the kindergarten sample: Omri, a second grader, lost only five teeth so far, the least among the three experiment students, and therefore the value three, which is close to five, will probably not be found among kindergartners [*366*]. Irit on the other hand claimed that the three is a possible outcome since it is close to two, the number of teeth Omri and herself lost in kindergarten [*367*].

Thus, when the students had to make a preliminary conjecture about outcomes of an unknown sample, they strongly built on their personal experiences (the number of teeth that they or their siblings lost) and their intuitive knowledge about the correlation between age and number of falling teeth, i.e., the older you get the more milk teeth you lose. This last piece of information was frequently used by them to support their reasoning.

When the three students arrived at the kindergarten, they found out that the kids were accidentally out for a field trip. The researcher therefore switched the data collection task to third grade. She seized the opportunity to again ask the students to conjecture on expected outcomes before they actually collected third grade data. After a short hesitation, the students suggested 12 or 13 teeth as probable outcomes, but later changed their hypothesis to "*seven and up*" range. This assumed range may be

[2] The kindergarten is part of the Kibbutz education system and is located in proximity to the primary school in which the experiment took place.

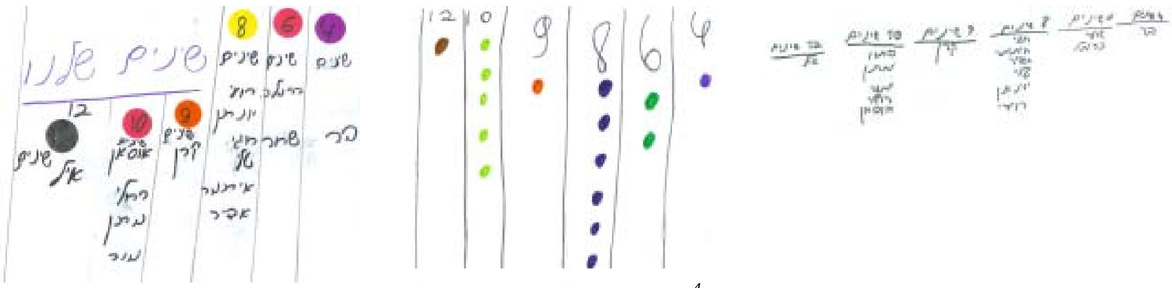
[3] All the transcripts are translated from Hebrew, and therefore may not preserve the original flavor of the language. [Square parentheses are used to complete missing words and add explanations that have been verified in the analysis process.]

related to the maximum value among the three of them (Yonat lost seven teeth). They seemed to again use their two main sources of knowledge, personal experience and the age-teeth relation. This correlation was explicitly articulated after they had completed their first survey among third graders. In response to the researcher's request to conjecture on second grade data, Irit definitely suggested that "*They* [second grade values] *are less* [than third grade]", and explained: "*Since we are younger and also our teeth grow slower, Ah … I mean falling slower*" [*1:594*].

Another unresolved issue at this stage was the role of *zero* as a legitimate data value. At this stage, zero is absent in their kindergarten conjectures. When the issue is discussed, they suggest to purposely ignore zero as an outcome, even if it was found in their survey. They were probably using their arithmetical knowledge that the zero is a neutral element in addition, and if the purpose is to sum up the numbers, it is useless and can be discarded.

5.3 FIRST "GROWING A SAMPLE" ACTIVITY: WHAT IF YOU MEET ADDITIONAL SIX THIRD GRADERS?

In the second session students discussed the results of their first survey, amalgamated their raw data and cleaned duplicate values, and individually created representations of the third grade data (see Figures 2 a-c).



a. Yonat *b. Irit (second inscription[4])* *c. Omri*

Figure 2: Students' inscriptions of falling teeth in the third grade sample (n= 16)

After long deliberations about these representations (which will be analyzed and reported elsewhere), the researcher has decided to engage the students in the first "growing a sample" activity: *"I am giving you pieces of paper - to write down a reasonable conjecture. The conjecture is: If you meet additional six third graders, how many falling teeth will they tell you? Each one of you will write down six numbers, six data values."*

The students started working on the task by writing down the title "Conjecture" on the top of their empty page and then prepared a column with six ordinal numbers (see Figure 3). Before they went on to conjecture, the following dialogue took place:

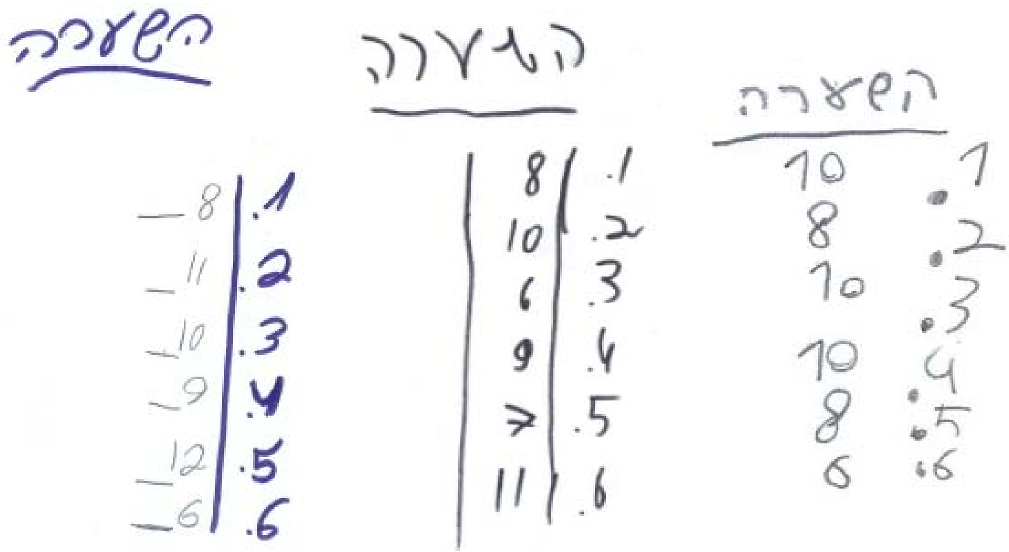
713	*Yonat*	But, it seems to me that I know all the numbers [in third grade]. [The video camera focuses on Omri writing down his conjecture, see Figure 3c.]
714	*Researcher*	OK, you can then repeat [numbers that you already know]. Is it OK to repeat numbers?
715	*Irit*	Sure.
716	*Researcher*	Ahah, why?
717	*Omri*	Because it is impossible that someone will tell me 16 [teeth].
718	*Researcher*	Ahah.
719	*Yonat*	If many kids had 16 [teeth]…
720	*Researcher*	What did many kids have?
721	*Yonat*	Eight teeth [the mode of the third grade sample].

(*Second session, 713-721*.)

4 Irit voluntarily created three different inscriptions for these data, but the discussion of them is beyond the scope of the current report.

(To view this discussion, see [Video Segment 2](#): *First 'growing a sample' activity: What if you meet additional six third graders?*)

Yonat was having difficulties to conjecture additional third grade outcomes since she was not sure about the task's purpose and the appropriateness of repeating values she already collected. She claimed that she knew all possible values for the third grade distribution. The researcher as well as her fellow students, Irit and Omri, legitimized repeating values, "*OK, you can then repeat*" [*714*]. Omri uses an interesting argument to support repetitions[5] in the conjecture: By pointing at an unusually large value for third graders [16 teeth], Omri explained [*717*] that the small range of probable values imposes repetitions in this task. The researcher uses Omri's comment to focus students' attention to a more probable value in the distribution, i.e., the mode eight. This is probably why Yonat and Irit started their lists with this value (Figure 3).



a. Yonat *b. Irit* *c. Omri*

Figure 3: Growing a sample conjectures for third grade. The column include ordinal numbers and the hypothesizes values.

After the students completed the conjecture task they were asked to reflect on their lists (Figure 3) and explain their particular choices. The following discussion took place:

726	*Researcher*	[Looking at Omri's conjecture, Figure 3c.] Omri, why did you choose these particular numbers, for example, 10 and 8? [Yonat is closely watching the discussion while Irit is busy with finishing up her conjecture.]
727	*Omri*	There are many kids that lost these numbers [of teeth].
728	*Researcher*	Ahah.
729	*Omri*	Look also here! [Points at the 8 and 10 the longest columns in Figure 2c, his representation of the third grade real sample.]
734	*Researcher*	[Yonat is handing her conjecture (Figure 3a) to the researcher.] OK, Yonat, what did you choose? [Researcher is showing Yonat's conjecture (Figure 3a) to everyone.]
735	*Yonat*	8, 10.
736	*Researcher*	Why did you choose these particular [numbers]? Why did you actually choose 8?
737	*Yonat*	8 is also a lot [Pointing at the 8 column in Figure 2a, her representation

[5] This is one possible interpretations of Omri's argument in row 717.

		of falling teeth in the third grade real sample.]
738	*Researcher*	And … Ahah
739 -741	*Yonat*	And 9 [another value in Yonat's conjecture, Figure 3a] … We did not investigate two classes, we investigated only one class. [Irit is loudly coughing in the background.] Therefore, it is not possible that only one kid, Keren from grade 3A, has lost 9 teeth. We did not investigate grade 3B, so we don't know. Therefore, I assumed 9 [teeth].
742	*Researcher*	I understand. Is it possible that we shall have 11 [another value in Yonat's conjecture, Figure 3a], which is a number that we didn't have so far [in the real data]?
743	*Irit*	Yes.
744	*Yonat*	Maybe … Yes, since we had 12 – then perhaps we'll have [11].
749	*Researcher*	[Researcher is introducing Irit's conjecture (Figure 3b) to the group.] No one has introduced 4 [minimal value in the real distribution]. Why not?
750 - 754	*Omri*	It is very small … Most of the kids don't have 4 … Bar [name of a third grader] is the only one that has [lost 4 teeth].

(*Second session, 726-754*)
(To view this discussion, see *Video Segment 3: Reflecting on growing a sample activity of third grade data*.)

The students used various considerations when they constructed their conjectures: repeated the frequent values, such as 8 (the mode) and 10 teeth; adding values with low frequency within the range of the real distribution (4-12 teeth), such as 9 teeth; avoiding values out of that range, such as 16 teeth; and adding missing values if they lie within the real distribution, such as 11 and 7 teeth. Yonat even implicitly mentioned the effect of increasing sample size on the probability to get some less frequent values [*739*]. They did not include extreme small case such as 4 teeth (nor 5 teeth), the minimal value of the real distribution. In rejecting this small value they probably used their background knowledge that four teeth is smaller than common values among the second graders. Here, and in many other cases during the experiment, they relied in their argumentation on contextual knowledge, that kids in third grade should lose more teeth than kids in their age group (second grade) and therefore values larger than 7 seem more probable than smaller ones.

It is interesting to point out that Omri was the only one that demonstrated a *distribution sense* with a center around 8–10 teeth, while the others have constructed a "*flat distribution*" with all its values appearing only once. By *distribution sense*, we mean an appreciation and understanding that a distribution of a variable tells us what values it takes and how often it takes these values. In other words acknowledging the two basic ingredients of a distribution: sample space and density (probability of frequencies). In contrast, the "*flat distribution*" students created hypothetical distributions by stating almost all probable values (the sample space of the variable), but did not provide any indication for their probable frequencies. They even included values that did not appear in the real data set as long as they were within a certain sensible interval for the specific age group, rejected outliers, and provided explanations for their choices. In the "*flat distribution*" perception, a new value in a sample space is considered as an important discovery which is highly valued. We can also coin this type of perception "*uni-dimensional perception*" of a distribution (versus a "*bi-dimensional perception*"); since it focuses attention just on the probable values in a distribution, and does not refer to frequency (the second dimension). This simple perception of a distribution can possibly clarify why Yonat did not see the point in repeating the same values she already knew [*713*]: the values are important, but not their frequencies.

The following graphs were created by the researchers to compare the three conjectured distributions and the real distribution, and to graphically demonstrate these suggested notions of students' perceptions (see Figure 4).

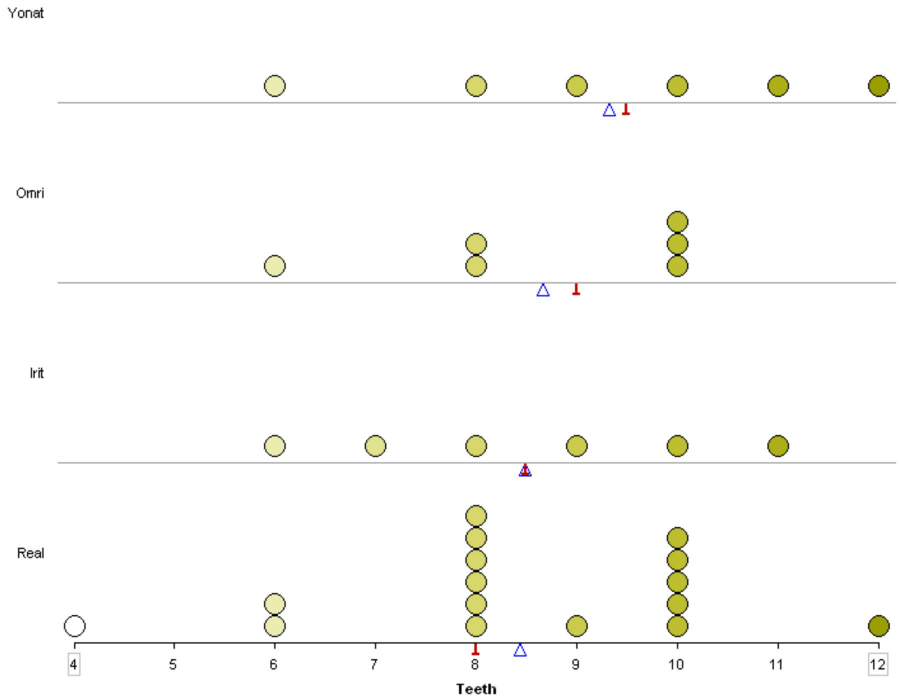


Figure 4: Students conjectures about a growing third grade sample compared to real data. The symbols and show the location of the mean and the median respectively.

Figure 4 better illustrates that students' conjectures were reasonable in terms of center and spread, slightly increasing the averages and decreasing the spread. All of them included values that had a frequency bigger than one (6, 8, and 10 teeth), and avoided the minimal value of four. However, Omri was the only one that demonstrated a *distribution sense* by creating a sensible left-skewed distribution with a "modal clump" around 8 – 10 teeth. The other two students focused on including almost all the "kinds" of values that are probable for third graders (sample space), which we coined as *flat distribution* perception.

This perception is also evident in the following discussion, which took place at the beginning of the third session. The researcher and Irit were reflecting in a personal interview setting on Irit's conjectured distribution created in the previous session (Figure 3b) as well as on her inscriptions of the real data (Figure 2b and two other graphs she made).

45	*Researcher*	[Irit's four inscriptions are placed on the table in front of her.] Why did you actually choose these numbers [in your conjecture, Figure 3b]?
48	*Irit*	Because many kids [third graders] can have 11, and 8, and 10.
49	*Researcher*	How do you know?
50	*Irit*	They [the third graders] are bigger [than us], and …
51	*Researcher*	You claim that many kids lost 11, 8 and 10 [teeth]. How do you know that? According to what do you know that?
52-58	*Irit*	According to the first list that I made [of real data, Figure 2b]. I have made here a list and it becomes clear that eight has the most.
59	*Researcher*	Ahah. And what else? You explained here the eight. Why did you choose 11, for example?
60	*Irit*	Because someone [in third grade] may have 11.
61	*Researcher*	Is it possible that we don't have even one 11 here [points at Irit's frequency graph of the real data, Figure 2b], and nevertheless there will be an 11 [in the conjecture]? [Researcher points at Irit's conjecture, Figure 3b.]
	Irit	[nodding in agreement]

	Researcher	Ok. Is it possible that … say I … shall choose 9 twice? Is it possible that a student will choose a number more than once in a conjecture?
62	*Irit*	It is possible, but it doesn't make sense.
63	*Researcher*	Doesn't make sense?
64	*Irit*	It is possible, but …
65	*Researcher*	Why? Please explain.
66	*Irit*	Since if presumably you'll be told once, and you will not find another one, then… If this will help you [to predict real values[6]].
67	*Researcher*	Ahah. You see here [points at Omri's conjecture, Figure 3c] the conjectures that Omri conjectured 10, 8, 10, 10, 8, and 6. What do you say about it?
68-72	*Irit*	It's not that good.
73	*Researcher*	Why?
74	*Irit*	Well … He [Omri] says that many [kids] will lose like this [8 and 10 teeth] and one will lose like this [6 teeth].
75	*Researcher*	Ahah. This is not good?
76-78	*Irit*	It's necessary to conjecture something else for each one [of the six numbers].
79	*Researcher*	Ahah. When you surveyed [third grade], each one was different from the other?
80	*Irit*	Almost[7].
81	*Researcher*	Yes? OK. I want to ask you a different question. How come there are many kids that lost 8 [teeth] and many kids that lost 10 [teeth], and only one child that lost 9 [teeth]?
82	*Irit*	Because there are [kids] that are bigger, and there are that are smaller, and the rate of their [loosing] teeth is not exactly the same.
83	*Researcher*	OK. And why there is no child that lost 7 teeth? I see here [Figure 2b] 6 and 8, How can that be?
84	*Irit*	I don't know.
85	*Researcher*	But do you think that if you go and ask more kids, they will then loose 7 teeth?
86	*Irit*	Yes.

(*Third session: 45-86*)

(To view this discussion, see *Video Segment 4: The emergence of the flat distribution idea*.)

One of the challenging elements in interpreting the above dialogue is the apparent difference between Irit's ability to account for the characteristics of an existing distribution and her difficulty to create a sensible hypothetical distribution. On the one hand, she easily remembered and identified the more frequent values and the mode [*48*, *52*], added probable values which were missing [*60-61*, *86*], and provided sensible explanations regarding the variability in the data [*82*]. But on the other hand, she refused to add frequencies to her conjectured values [*62*] and criticized Omri's choice of a hill-shaped distribution [*68*] since "*It's necessary to conjecture something else for each one*" [*76*]. One possible explanation to this conceptual gap between the two actions can be related to the small sample size of the conjecture (n=6), which does not impose on students to deal with frequencies.

In the third session, before they collect data from their second grade peers, students were asked again to conjecture the values of six imaginary second graders. Similar processes and considerations were repeated by the students, except for Omri's hypothetical distribution which becomes "flat" and contains no indication of frequency (see Figure 5 – an illustration of their hypothetical distributions compared to the real distribution which was later collected by them). When Omri was asked to explain why he did not repeat any value, he reported that he had known that third graders had repeated values, and therefore should have had to take that into consideration. We assume that he did not do that,

6 This is a possible interpretation, but we are not certain what she meant.

7 This is a surprising response, since the sample data included many identical values, such as six students lost eight teeth (Figure 2b). An alternative interpretation, that can resolve this contradiction, is that Irit was referring to the values of the sample space. Thus, she was focusing on the fact that almost all the space interval (4 to 12) was covered by third graders' responses, i.e., 4, 6, 8, 9, 10 and 12.

because he was not certain about the trend in the second grade data, (e.g., which values are more frequent) as he was in the previous situation, having collected a sample of third graders before conjecturing.

The students asserted that if they gave room to frequency in their conjectures they would not be able to fully cover the spectrum of possible values. In their view, a hypothesized distribution was *representative* or *typical* of a real situation if and only if it covered most of its sample space. Thus, we reiterate our distinction between primitive distributional conception, which has one dimension (*uni-dimensional perspective*) and a more developed conception that includes both values and their frequencies (*bi-dimensional perspective*).

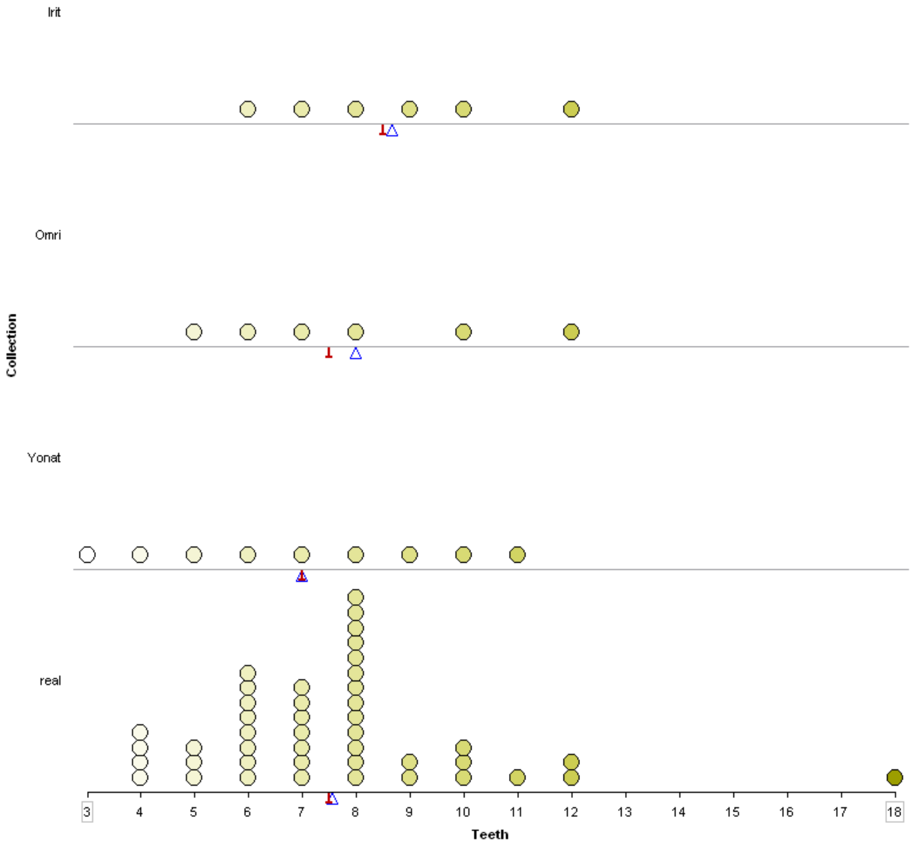
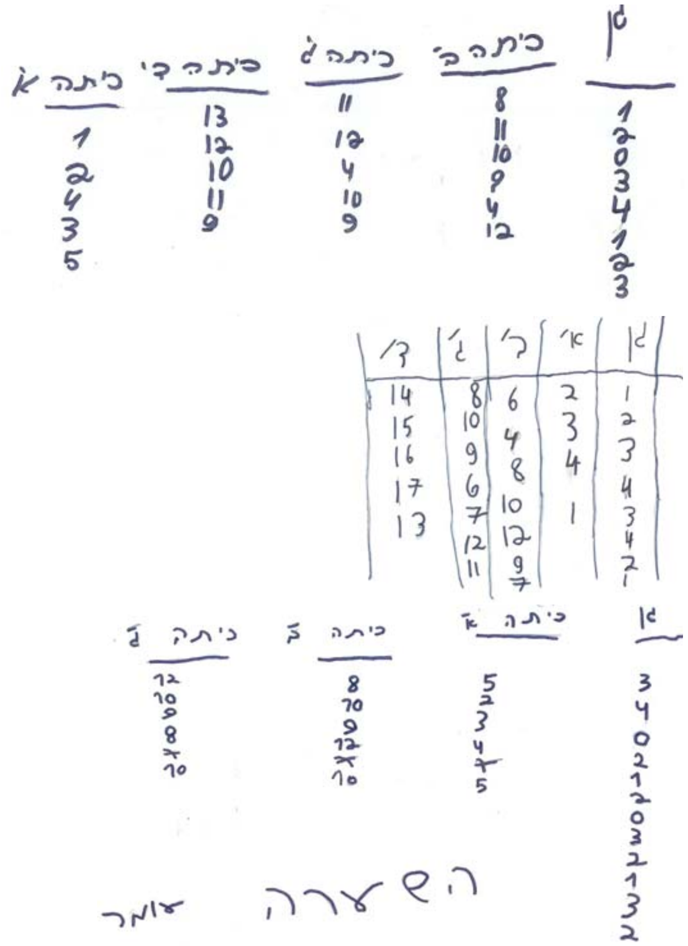


Figure 5: Students conjectures about second graders compared to real data that was collected afterwards. Although the researcher asked for six values, Yonat offered 9 values.

5.4 COMPARING HYPOTHETICAL DISTRIBUTIONS

In the fourth session, students were asked to create several hypothetical distributions and compare between them, to allow us better understand their actions and considerations while reasoning about distributions. Students were requested to produce hypothetical data sets for grades K-3, with 6 values each. At this stage, only grades 2 and 3 were actually surveyed by the students. Although each one of them worked separately, they produced a similar tabular structure in which each column stood for a different grade, starting from Kindergarten on the right (see Figure 6).

Yonat and Irit first produced only four values for kindergarten and first grade. Even when prompted directly by the researcher to add more values they insisted that it is impossible, since "*They* [the kindergartners] *will soon become second graders*" (4:248), or: "*The* [additional] *numbers of teeth will belong to second grade*" (4:251). We see these responses as an indication of their persistent "*uni-dimensional*" view of distributions in conjecture tasks, focusing on sample space and not on frequencies. Only the researcher's insistence on additional values forced them to repeat some of the values in their conjectures. Omri, on the other hand, produced all his hypothetical distributions with a single mode by repeating one value, which he seemed to consider more probable. We see that as an indication of his emerging *bi-dimensional* perception of distribution. Omri is also able to explain his choice, for example: "*In the kindergarten, a kid does not lose, say, five teeth; neither four nor ten. They lose fewer than four and that's it … and they are 30 kids* [in the kindergarten], *that's many*" (4: 267-271).



a. Yonat

b. Irit

c. Omri

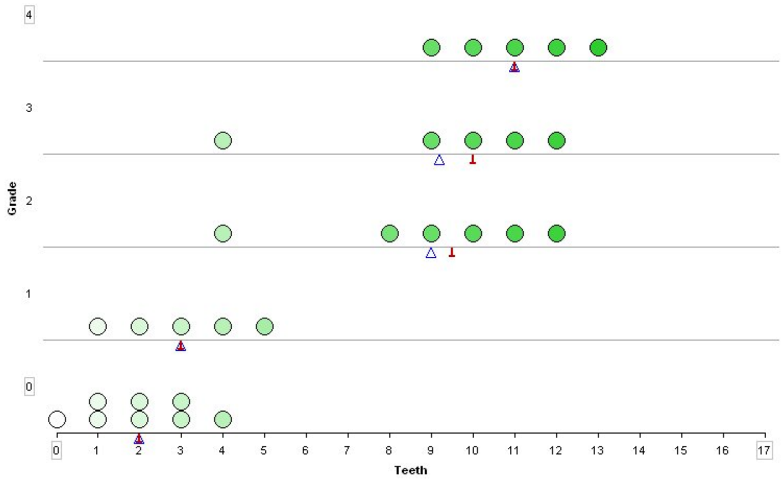
Figure 6: Stidents constructing hypothetical distributions of grades K-4

Figure 7 was created by us to graphically compare the students' hypothetical distributions. All students displayed an awareness of the growing trend of the phenomena, i.e., the positive relation between grade level and the number falling teeth. Furthermore, they seemed to have had a strong sense of the probable values at each grade level. Two of them mentioned for example that they expected that the mode in the kindergarten sample would be two teeth ("*would repeat itself more than anything else*"). The fact that these hypothetical distributions were produced in a table formant (versus a more useful graphical inscription), and that the values were introduced in an unordered fashion, reinforces our impression regarding their strong contextual background knowledge. Still, students are not consistent in describing the growing trend with age, for example, two consecutive grades are identical (Grades 2 and 3 in Omri's conjecture, Figure 7c) or almost identical (Grades K and 1 in Irit's conjecture have the same range, Figure 7b).

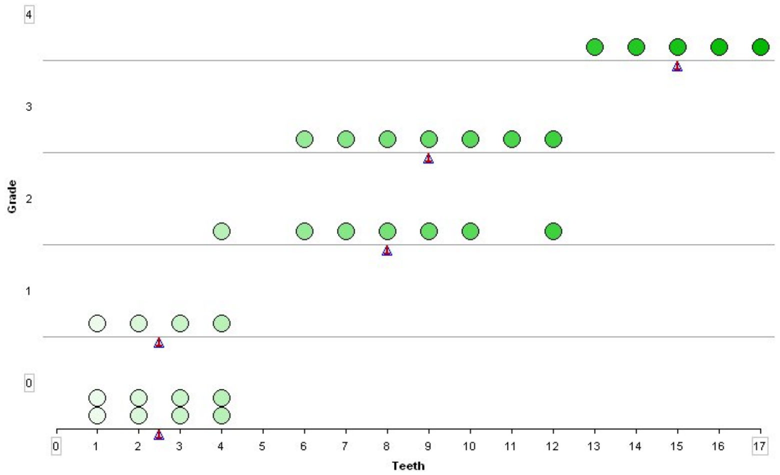
6. DISCUSSION

This study explored the emergence of second graders' informal reasoning about distribution in a carefully planned learning environment that included extended encounters with open-ended Exploratory Data Analysis (EDA) activities. We carefully examined the process of constructing meanings, language, representations and appreciation for distributions and concentrated on the detailed qualitative analysis of the ways by which three second grade students (age 7) started to develop views (and tools to support them) of distributions in investigating real data, inventing and using various informal data ideas and representations.

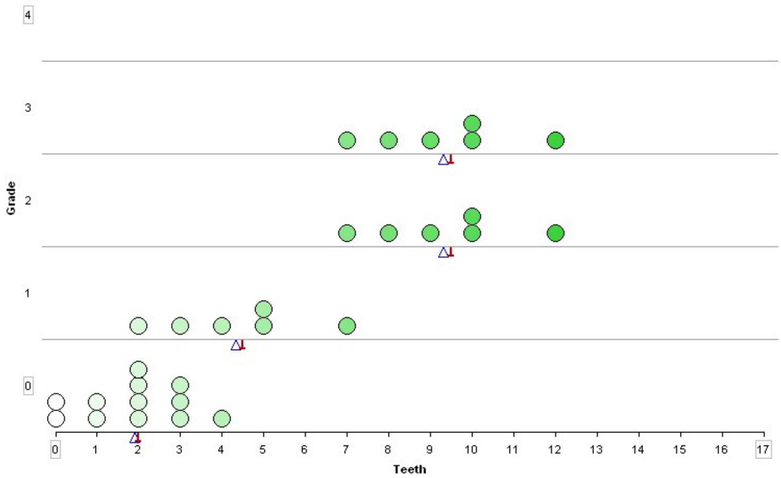
The analysis of data sheds light on some of the simplest emerging conceptions and representations of distributions that students invent, construct, informally recognize and can understand. In particular interest is the distinction between the "*flat distribution*" conception and the emerging *distribution sense*. Students with "*flat distribution*" conception created hypothetical distributions by stating almost all probable values of the distribution (the sample space), but did not provide any indication for their probable frequencies. Students can include values that did not appear in the real data set that was previously collected, but were within a certain sensible interval for the specific age group. They tend to reject outliers and can provide explanations of their choices. In this perception a new value in a sample space is considered as an important discovery which is highly valued. It seems that this simple distributional conception can become the runway to the emerging *Distribution sense*. This perception includes an appreciation and understanding that a distribution of a variable tells us what values it takes and how often it takes these values. In other words acknowledging the two basic ingredients of a distribution: sample space and density (probability of frequencies). We also coined this distinction with the notions of "*uni-dimensional perception*" of a distribution versus a "*bi-dimensional perception*". There is some evidence that older students also handle the complexity of distributional reasoning by reducing complexity, e.g., focusing on partial features of a distribution (e.g., Moritz, 2004; delMas & Liu, 2005).



a. Yonat



b. Irit



c. Omri

Figure 7: A representation of students' comparing hypothetical distributions of grades K-4.

How the transition between these two distributional conceptions is made, is a matter of interest to researchers and statistics educators. It is not clear whether the *flat distribution* perception is a result of cognitive overload, a developmental stage of noticing only one element (out of the many distributional characteristics), or an issue of maturity. Further study in this direction will be of great importance to direct our instructional efforts at an early age. Furthermore, it is not clear whether the *distributional sense* that began to emerge in students' responses is a link in a string that will evolve at older age students to more complex, aggregate-based understanding of distributions with shape, center and spread, and later an abstract conceptual entity (theoretical distribution and sampling distribution).

The three students described in this study were considered by their teacher to be both able and verbal. Their choice was aimed to enable the collection and analysis of focused and remarkably detailed data in order to draw, in very fine strokes, the "picture" of their emerging statistical reasoning about distribution. Even when a phenomenon seems important and the data interpretation was validated and agreed upon, the question of the idiosyncrasy of the identified phenomenon may remain open. Therefore, in future studies, the data and interpretations from students in other second grade classes will assist in checking for generalizability of the phenomena.

7. IMPLICATIONS

The idiosyncratic aspects of this study restrict the provision of broad recommendations. However, several conclusions that are tied to specifics of this study and its results, in the context of results from similar studies, can be drawn. The learning processes described in this paper took place in a carefully designed environment. This environment included: a learning trajectory built on the basis of expert views of EDA as a sequence of semi-structured (yet open) leading questions within the context of extended meaningful problem situations (Ben-Zvi & Arcavi, 1998), timely and non-directive interventions by the researcher as representative of the discipline in the classroom (cf., Voigt, 1995), and several repetitions of tasks to enable students handle complex actions (constructing inscriptions of data, conjecturing, growing a sample, etc.) leaving time and energy for conceptual discussions.

In learning environments of this kind, from the very beginning students encounter, develop, and work with ideas, concepts, cognitive tools and dispositions related to the culture of EDA, such as making hypotheses, summarizing data, recognizing trends and variability, identifying interesting phenomena, comparing distributions and handling data representations. Skills, procedures and strategies, such as creating and interpreting graphs and tables, are learned as integrated in the context and at the service of the main ideas of EDA.

It can be expected that beginning students will have difficulties (of the type described) when confronting such problem situations. However, it is proposed that what these students experienced should be an integral and inevitable component of a meaningful learning process if it is to have lasting effects. If students were to work in environments such as the above, teachers are likely to encounter the following learning phenomena:

- Students' prior knowledge would and should be engaged in interesting and surprising ways – possibly hindering progress in some instances but making the basis for construction of new knowledge in others (e.g., the *total perception* of a distribution).
- Many questions that would either make little sense to the students, or, alternatively, will be re-interpreted and answered in different ways than intended.
- Students' work that would inevitably be based on partial understandings, which will grow and evolve.

This study suggests that in order to help students gradually build a sense of the meaning of the data and statistical task with which they engage, multiple factors can and should be planned. These include appropriate teacher guidance, peer work and interactions, and more importantly, ongoing cycles of experiences with realistic problem situations.

Given that it is difficult to tease out the effects of what students learned or could or couldn't do from the support of the researcher, further study is recommended that focus more attention on the role of teachers and what they should do, or learn to do, in order to promote statistical reasoning about distribution. Much of students' progress in the current study is influenced by their interactions with the researcher that helped them adopt the statistical perspective but did not instruct them in exactly what to do or how to reason. The role of the teacher deserves further exploration.

It is generally recommended that students be provided with multiple opportunities to engage with data in group-comparison tasks. The students in this study have gained from reading and interpreting multiple types of self created data inscriptions. The role of student-invented data representations and new graphical tools available through educational software and Internet has to be investigated to better expose the many ways distribution is noticed and understood by students. It is hoped that the complexity involved in group-comparison tasks can push students to think about the meaning of what they do and how they reason in statistics, develop relevant actions and interpretations, and be more critical of their actions and interpretations.

REFERENCES

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal*, *3*(2), 64–83.

Bakker, A., & Gravemeijer, K. (2004). Learning to Reason about Distribution. In D. Ben-Zvi, & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 147–168). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ben-Zvi, D. (2002). Seventh grade students' sense making of data and data representations. In B. Phillips (Eds.), *Proceedings of the Sixth International Conference on Teaching of Statistics* (on CD-ROM). Voorburg, The Netherlands: International Statistical Institute.

Ben-Zvi, D. (2004). Reasoning about Data Analysis. In D. Ben-Zvi, & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 121–146). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ben-Zvi, D., & Arcavi, A. (1998). Toward a characterization and understanding of students' learning in an interactive statistics environment. In L. Pereira-Mendoza (Ed.), Proceedings of *the Fifth International Conference on Teaching Statistics* (Vol. 2, pp. 647–653). Voorburg, The Netherlands: International Statistical Institute.

Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics*, *45*, 35–65.

Ben-Zvi, D., Garfield, J., & Zieffler, A. (in press). Research in the Statistics Classroom: Learning from Teaching Experiments. In G. Burrill, P. C. Elliott, F. Bullard, A. Flores, S. A. Hess, J. Landwehr, and C. Pateracki (editors), *Thinking and Reasoning with Data and Chance* (The 2006 NCTM Yearbook). Reston, VA: National Council of Teachers of Mathematics.

Bright, G. W. & Friel, S. N. (1998). Graphical representations: Helping students interpret data. In S. P. Lajoie (Eds.), *Reflections on statistics: Learning, teaching, and assessment in grades K-12* (pp. 63–88). Mahwah, NJ: Lawrence Erlbaum.

Cobb, G., & Moore, D. (1997). Mathematics, statistics and teaching. *The American Mathematical Monthly*, *104*(9), 801–823.

Cobb, P. (1999). Individual and Collective Mathematical Development: The Case of Statistical Data Analysis. *Mathematical Thinking and Learning*, *1*(1), 5–43.

Cobb, P., & McClain, K. (2004). Proposed design principles for the teaching and learning of elementary statistics. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 375–396). Dordrecht, The Netherlands: Kluwer Academic Publishers.

delMas, B., & Liu, Y. (2005). Exploring Students' Conceptions of the Standard Deviation. *Statistics Education Research Journal*, *4*(1), 55–82. Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ4(1)_delMas_Liu.pdf

DfES. (2000). *Mathematics: The National Curriculum for England*. London: Department for Education and Employment.

Druin, A. (1999). Cooperative inquiry: Developing new technologies for children with children. Proceedings of Conference on Human Factors in Computing Systems (CHI) *99*, 592–599. New York, NY: *Association for Computing Machinery (ACM).*

Edelson, D.C. (2002) Design research: What we learn when we engage in design, *The Journal of the Learning Sciences*, *11*(1), 105–121.

Freudenthal, H. (1991). *Revisiting Mathematics Education: China lectures*. Dordrecht: Kluwer Academic Publishers.

Gal, I., Rothschild, K., & Wagner, D. A. (1990). *Statistical concepts and statistical reasoning in school children: Convergence or divergence?* Paper presented at the annual session of the American Educational Research Association, Boston.

Garfield, J. B., & Ben-Zvi, D. (in review). *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice*. Emeryville, CA: Key College Publishing.

Gravemeijer, K. P. E. (1994). *Developing Realistic Mathematics Education*. Utrecht: CD-ß Press / Freudenthal Institute.

Gravemeijer, K. P. E. (2000). A rationale for an instructional sequence for analyzing one- and two-dimensional data sets. Paper presented at the annual meeting of *the American Educational Research Association*, Montreal.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27*(3), 337–364.

Konold, C. (2002). Teaching concepts rather than conventions. *New England Journal of Mathematics, 34*(2), 69–81.

Konold, C., & Higgins, T. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin, & D. E. Schifter (Eds.), *A Research Companion to Principles and Standards for School Mathematics* (pp. 193–215). Reston, VA: National Council of Teachers of Mathematics.

Konold, C., Pollatsek, A., Well, A., & Gagnon, A. (1997). Students analyzing data: Research of critical barriers. In J. B. Garfield & G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics* (pp. 151–167). Voorburg, The Netherlands: International Statistical Institute.

Konold, C., & Pollatsek, A. (2002). Data Analysis as the Search for Signals in Noisy Processes, *Journal for Research in Mathematics Education*, *33*(4), 259–289.

McClain, K., & Cobb, P. (2001). Supporting students' ability to reason about data. *Educational Studies in Mathematics, 45*(1-3), 103-129.

Meira, L. R. (1991). *Explorations of Mathematical Sense-Making: An Activity-Oriented View of Children's Use and Design of Material Displays*, unpublished Ph.D. dissertation, Berkeley, CA: University of California.

Meletiou, M. (2002). Conceptions of variation: A literature review. *Statistics Education Research Journal, 1*(1), 46–52.

Mokros, J., & Russell, S. J. (1995). Children's Concepts of Average and Representativeness. *Journal for Research in Mathematics Education, 26*(1), 20–39.

Moritz, J. (2004). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 227–256). Dordrecht, The Netherlands: Kluwer Academic Publishers.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Petrosino, L., Lehrer, R., & Shauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical thinking and learning*, 5(2&3), 131–156.

Piaget, J. (1971). *Psychology and Epistemology: Towards a theory of knowledge.* New York: Viking Press.

Piaget, J. (1973). *To understand is to invent: The future of education.* New York: Grossman.

Reading, C., & Shaughnessy, M. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 201–226). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schoenfeld, A. H. (1994). Some notes on the enterprise (research in collegiate mathematics education, that is), *Conference Board of the Mathematical Sciences Issues in Mathematics Education*, *4*, 1–19.

Siegler, R. S., & Crowley, K. (1991). The Microgenetic Method: A Direct Means for Studying Cognitive Development. *American Psychologist 46*(6), 606–620.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education*. Dordrecht, The Netherlands: Kluwer Academic Publisher.

Voigt, J. (1995). Thematic patterns of interaction and sociomathematical norms. In P. Cobb & H. Bauersfeld (Eds.), *Emergence of mathematical meaning: Interaction in classroom cultures*, 163-201. Hillsdale, NJ: Erlbaum.

Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. M. Cole, V. John-Steiner, S. Scribner & E. Souberman (Eds.), Cambridge, MA: Harvard University Press.

Watson, J. M. (2004). Developing reasoning about samples. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 277–294). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Watson, J. M., & Moritz, J. B. (1999). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics 37*(2), 145–168.

Wild, C. (2005). Key address. SRTL-4 Proceedings. [CD-ROM]

Zawojewski, J. S., & Shaughnessy, J. M. (2000). Mean and median: Are they really so easy? *Mathematics Teaching in the Middle School*, *5*(7), 436–440.

ANNOTATED LIST OF VIDEO SEGMENTS

The following video segments are hyperlinked to this paper.

#	**Title**	**Length** (Min:Sec)
Video Segment 1	First conjecture: How many teeth do kindergartners and third graders lose?	00:58
Video Segment 2	First 'growing a sample' activity: What if you meet additional six third graders?	00:46
Video Segment 3	Reflecting on growing a sample activity of third grade data.	01:35
Video Segment 4	The emergence of the flat distribution idea	03:50

DANI BEN-ZVI
Faculty of Education, University of Haifa
Mt. Carmel, Haifa, 31905
Israel