Interplay of the Number Concept and Statistics Using Attitudes and Techniques of EDA

Andrejs Dunkels - Luleå, Sweden

1. Introduction : the number concept

"If mathematics is just a sequence of recreational scratchings, then why do the games we engage in prove so useful?" (Philip Kitcher, 1984)

Most syllabuses of today would, in some way or another, include a statement that mathematics for the elementary school must deal with the real world. One of the tasks of mathematics is to provide a variety of tools for that. Mankind has developed a number concept through practical applications, through the need to possess suitable ways of describing phenomena of life, of describing phenomena of the world around us - from trivial day-to-day matters to sophisticated branches of science, technology, social studies, medicine, economy. To be brief, wherever some process of thought is involved there is a need for descriptions, many of which tend to be quantitative.

Much of the mathematics of the elementary grades is in fact part of what is traditionally called statistics, since much of it could be termed description of the world around us. So statistics is an integral part of the teaching and learning about numbers and their properties. It is rewarding to try to replay, or at least to some extent imitate, the course of history; the learner benefits from going through some of the stages that our ancestors have had to pass. The concept of number has developed in part due to curiosity, in part due to practical needs in the desire to increase the quality of life.

2. A fruitful example

"The individual is not alone, but is part of a society. Reality is also social. The interplay of the natural and the artificial in building up environmental reality is probably one of the most critical areas in which mathematics education has a major role to play." (Ubiratan D'Ambrosio, 1986)

2.1 *Pre-schoolers*

In Sweden children begin school at the age of seven. Furthermore, pre-school has been made available to all six-year-olds.

In most pre-schools the children bring their own fruit each day. In one such group my colleague Birgitta Wallström found that the children were always curious about how many had this fruit and how many had that, and, in particular, how many have the same as I. This called for systematisation, the need was clearly felt for fruit statistics.

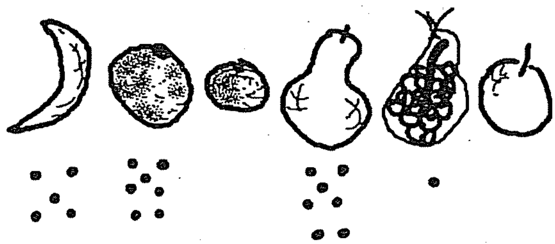
At first the children showed what they had by raising hands and counting. This was very exciting.

The first real diagram emerged on the floor. The children displayed in neat rows all the actual fruits they had brought that day. There was one row of bananas, one of oranges, one of mandarin oranges, one of pears, one row consisting of a plastic bag with grapes, and one of apples.

After a while some children felt uneasy about their fruit lying there, unprotected, on the floor. They wanted to have it back, and so there was a need for recording before the "diagram" disappeared.

A suggestion from the group was to draw a picture of each type of fruit and then to use one dot for each piece of fruit. For most children this was their first conscious matching experience. They established a one-to-one correspondence between the fruits and the dots. This is the basic idea of counting. One builds a reference set by matching each element of the set under consideration with a unique element of the reference set. In this way statistics is used as a means of building up the number concept.

Each day at least some of the children made the fruit diagram of that day. This is what it looked like (the picture is a reproduction, not a true copy).



The interest in the fruits of course was a golden opportunity for practicing counting, and getting to understand what counting is all about. Many questions arose quite naturally, pinpointing some of the problems of collecting real data:

- Are we to include the fruit I ate on my way here?
- I have lost an apple. Should I count it anyway?
- Shall I count the *children* who have brought grapes or all the grapes? Or perhaps the bunches.

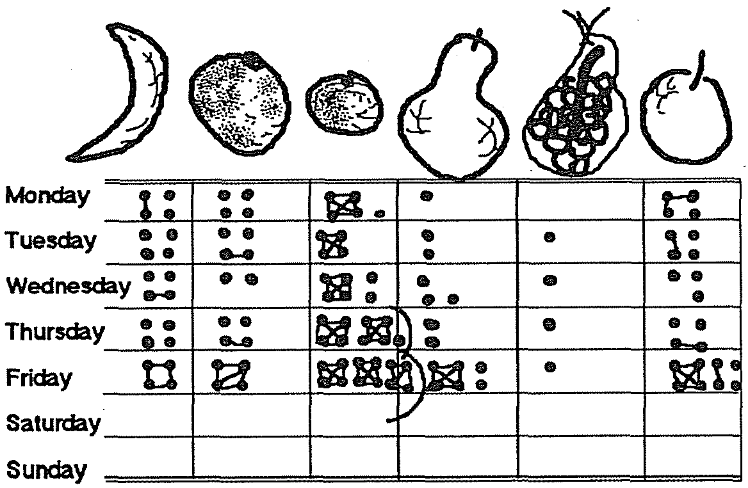
All the possibilities were tried and investigated, though different possibilities on different occasions. There was no reason to rush things.

As a by-product of all this the children got engaged in a statistical investigation of grape bunches. I have selected one of the findings:

- There may be 17 grapes to a bunch!

Fruits during a week: It is important for pre-schoolers to get involved in natural counting situations. They need to experience instances where handling of numbers occurs in a non-artificial way. Pre-schoolers often seem to know what numbers are about. They exhibit confidence and knowledge, for example, when buying sweets. However, this might be just shallow knowledge, with a base neither in concept mastering nor understanding. Neuman (1988) has said much of interest about such matters.

The children wanted to keep a record to see how the number of fruits varied during a whole week. Birgitta then introduced tallying by tens (Tukey, 1977), which was very well received. Some children redrew previous diagrams, and all were seriously engaged in producing the week chart of fruits. The next picture gives an idea about the result. The first such chart was made collectively on a huge paper sheet.



Then the children were given sheets that Birgitta had prepared in advance with the names of the days of one week written, and pictures drawn of those fruits that usually occurred. Later each child produced everything by himself or herself.

After having completed a display the time has come for the posing of questions. What does the display tell us?

- Which fruit was the most popular? How come?
- Was it always the same child who had grapes?
- How many of each fruit were there during the whole week?
- What happened on Friday? (*Answer*: A grade 2 class from the neighbouring school came for a visit, and the treat then was fruit salad, a choice inspired by the fruit statistics.)

3. Numbers and heights

"In the genesis of the number concept the counting number plays the first and most pregnant role." (Hans Freudenthal, 1973)

In Sweden the arithmetic of first-graders normally covers numbers below 100, and so towards the end of Grade 1, exercises focusing on tens and units are appropriate.

I have previously reported about such exercises in Grade 1 using stem-and-leaf displays (Dunkels, 1986; Pereira-Mendoza and Dunkels, 1989).

In order to possess a true number sense of natural numbers it is necessary to be able to perceive simultaneously the cardinal and ordinal aspect (Neuman, 1987; Kamii, 1985, 1990). This is something that grown-ups do automatically and unconsciously. A characteristic of a true number sense, or the possession of the number concept for natural numbers, is the ability to switch between these two aspects.

Thus the teaching of number should include exercises where both aspects appear. In most of the current textboks there are no such exercises. Again the stem-and-leaf display turns out to be a suitable tool.

I wanted to collect data from the children's own environment in a Grade 1 class. The pupils' heights seemed to be something that might be of interest to the children. Now all children would be more than 100cm tall, which is outside the number range covered by the majority. Therefore we made a mark on the wall 1m above the floor, and found the number of cm above that mark, in other words the excess over 1m of each child's height. The completed stem-and-leaf display is given below.

	Tens	cm	Ones	
1	2	|	1	(1)
(8)	2	|	66667889	(8)
6	3	|	033	(3)
3	3	|	77	(2)
1	4	|	2	(1)
				(15)

Before discussing the lesson let us look at the way numbers are used in the stem-and-leaf display. Following Tukey (1977) the number of digits of each row is given to the right. Brackets are used to avoid mixing digits up. Furthermore the brackets assist in keeping the functions of the various numbers apart. The numbers in the rows are part of the data under consideration. These numbers arose as the result of measuring, not of counting. This is often the case. The natural numbers in brackets are used as cardinals, they tell us how many data values there are in each row. Thus the stem-and-leaf display offers a concrete opportunity of counting and finding "how many" in passing, so to speak, unnoticed, while working with the data under consideration.

To the left are the cumulated frequencies from both ends, except for, as usual, the row containing the median. That left column is very interesting. It contains natural numbers. This time their function is not cardinal but ordinal, they function as counting numbers. The nature of counting is ordinal, and one way of obtaining the cardinal number of a finite set is to take the elements one by one, saying *one*, *two*, *three*, ... but in fact doing *first*, *second*, *third*, ..., and if the *n*:th element is the last, then the conclusion is that the cardinal of the set is *n*.

I do not mean to say that, when children encounter these displays for the time, we must necessarily explicitly write all the numbers I have included in the stem-and-leaf display. Not until the teacher feels that the class is ready for it do those two columns need to be introduced. The numbers are included here in order to facilitate the discussion of their use and function. However, even if we do not write the numbers of the left-most and right-most columns, then concepts such as "the number of data values of each row", "the smallest data value", "the second value counted from the smallest", "the middle value", etc.:, of course can still be used.

In teacher education it is important to shed light on the different aspects of number. I have used stem-and-leaf displays with student teachers since 1980 and been very pleased with the possibility of discussing these aspects in connection with something having to do with practical mathematics. When the cardinal and ordinal aspects appear in a natural context their meanings are highlighted more clearly than when using made-up examples, hence favourable conditions for learning are at hand.

One could express the situation thus: Tukey makes ingenious use of our system of writing numbers. He sees how the place value system can assist in making us see better what data tell us, and he also utilises the interchangability and simultaniety of the cardinal and ordinal aspects of natural numbers.

What I try to do then is to take advantage of all this and apply it the other way round. By having pupils identify numbers and their use they instead learn what the properties and aspects of numbers are.

As a teacher educator I am attached to a school where some of our student teachers do their teaching practice. I meet children about once a week.

When I met the class to discuss stem-and-leaf plots the class teacher acted as an observer. We discovered that the activity could be seen not only as describing a part of real life and a way of reviewing some properties of numbers, but also as a means of assessment or diagnosis. Firstly we could observe the state of the skill of the actual writing of numerals, without having to tell that we were observing that. Secondly we could judge how the pupils were able to make use of their knowledge of tens and ones. In particular the class teacher was pleased to get an extra opportunity of assessing this detail.

With data from real life there are often unforeseen problems. Not everybody wanted to participate in our activity. Two short pupils did not want to be measured, and of course they did not have to. However, they were happy to take part in the recording of the classmates' data, they were happy to discuss, ask, and answer questions.

The questions of earlier activities on other occasions could now be recycled with minor adjustments (Dunkels, 1986):

- What is the smallest amount over 1m in this class?
- The largest?
- What is the difference between the smallest and largest?
- Did we get all possible amounts in cm between the smallest and the largest?
- How many such possible amounts are there?
- Tell some that are missing.

As usual the children were encouraged to pose questions themselves:

- How many centimetres do I have to grow to be as tall as the tallest in our class?
- Are there any pupils who are the same length?
- How many?
- How many are 30cm or more? (With the tacit addition: taller than 1m.)
- When the shortest will be 30cm over 1m, how much will the tallest be over 1m?

The last question gave me the idea to ask what the display would look like after the summer holidays. We agreed that there was no way of knowing, the order between pupils might change, so I am really looking forward to meeting the class again in September. We will then make a display of the changes too.

4. Scales and rivers

"Children are immersed in a reality. But which reality?" (Ubiratan D'Ambrosio, 1984)

Like many educators I have given much thought to the question of what children's reality is like, and, in particular, what children would like to investigate at school that would be interesting.

Advertisements are a rich source in education. There is much material suitable for statistics, and so I have often asked pupils to bring their own advertisements to school for us to explore and engage in problem posing, advertisements they cared for. Allow me to mention that so far the outcome is disappointing. The pupils have never taken the task of finding advertisements seriously and just brought advertisements about special offers for flour, napkins, and other products, revealing the last-minute character of their data collection.

Would fifth-graders, which in Sweden means 11–12-year-olds, be interested in some matters connected with geography? Is it at all interesting to display lengths of rivers, heights of mountains, volcanos, caves? I recall that I became very interested in the example about the heights of 219 of the world's volcanos in Tukey (1977), and so I thought that perhaps some fifth-graders would feel the same about geographical data.

I was going to meet a Grade 5 class that I have met with at least a couple of times each term since they were in Grade 1. So I therefore know the pupils quite well. In Grade 1, incidentally, I worked in this class with their parents' ages, reported in the section *The first EDA lesson* in Dunkels (1986).

My objective was now to find out how the pupils could handle scales and the number line. They had so far not had any formal instruction about the number line, which, in Sweden, is usually introduced in Grade 6. I also wanted an opportunity to review stem-and-leaf displays.

I decided that we would look at the lengths of rivers. Luleå lies at the mouth of the big Lule river. Do the pupils have any idea of how long Lule river really is?

This was the first question. Each pupil wrote down his or her estimate. We reviewed the construction of a stem-and-leaf display with this data. We had decided to use the unit *mil* = 10 km, which is the most natural unit to be used by Swedes for this kind of length.

Each pupil of the 13 present then called out his or her guess, and I recorded it in the stem-and-leaf format. The result was, as usual, an unordered display. Each pupil was then supposed to construct the ordered display. Here is the completed display.

mil = 10km

	Tens	Ones	
2	1	00	(2)
6	2	0013	(4)
(2)	3	55	(2)
5	4	038	(3)
2	5		(0)
2	6	05	(2)
			(13)

The median is here 350km. The length is actually 450km. The pupils were next given a list of the 13 longest rivers of Sweden, in alphabetical order, and were instructed to make a stem-and-leaf display for the lengths of the rivers. The lengths were given in kilometres, but all the values ended in a zero, so the lengths could equally well be seen as given in the Swedish unit *mil*.

mil = 10km

Tens	Ones	
2	7	(1)
3	57	(2)
4	1233556	(7)
5	27	(2)
6		(0)
7	2	(1)
		(13)

The usual way of introducing the number line in Swedish schools is to start from the abstract, i.e. to start with numbers in a context-free setting. Applications come later. I wanted to do it the other way round. I wanted to make the pupils acquainted with the number line through its use in some context.

First of all, with a simple test, I had found that the pupils had surprisingly thorough knowledge of the plotting of points on a given scale and the reading of a scale. All the test items were related to contexts familiar to the pupils. Children of course acquire some knowledge, without formal instruction, of often-encountered phenomena of their society (D'Ambrosio, 1986). In that perspective it is perhaps not surprising that children in Sweden pick up features of the ways of communicating information we very often use, e.g. a line equipped with a scale.

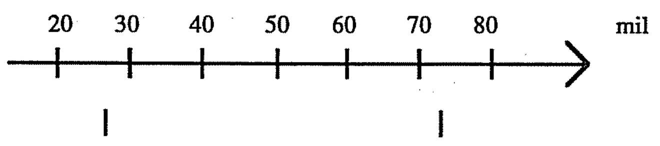
I had expected the introduction of box-plots to require some preliminaries about scales. This was not necessary.

What turned out to be difficult, however, was the choosing of the unit and the making of one's own scale, one of the issues discussed within the subject of descriptive statistics. This is a task where one has to be active, to use properties of numbers, and take one's own initiative, which requires energy and hence is difficult. One must also be prepared to re-do the scale at least a couple of times, which is hard for many learners.

As an intermediate step toward box-plots I had decided to use what I call the *mmm*-plot, which is obtained by stopping about halfway through the construction of a box-plot. One draws small line segments for just three of the distinguished data values, minimum–median–maximum. After having done that one stops. No fourths, no connecting lines, nothing more. The *mmm*-plot is ready.

For some pupils the box-plot requires so many steps that the handling of them all obscures the looking at the data. So the idea is to have a simpler plot to begin with. The *mmm*-plot is easy to draw, no pupil will be left behind, and there is much information to be obtained from *mmm*-plots too.

The *mmm*-plot was introduced to Class 5A by carrying on with the 13 longest rivers of Sweden. The pupils were introduced to the two extreme *m*:s, and were then asked to draw a line, select an appropriate scale, and then mark, with two short lines below the scale, the positions of the two extreme *m*:s, as shown below.



The next step to be taken was then to discuss what is meant by the *middle* value. It is interesting to note that the pupils showed that they knew about averages, in terms of the arithmetic mean. Many pupils suggested that as the value for the middle *m*. Of course that is one reasonable choice.

Are there other possibilities? The middle value could refer to the interval between the extreme *m*:s, so we could choose the midpoint of that interval as the middle value. That choice would give what Tukey calls the mid-range and what traditionally is called the mid-extreme.

I told the class that I had a very particular choice for the third *m* in mind, and gave them the hint that I wanted the *m* be such that there are as many of the data values to the left of it as there are to the right.

It did take some reasoning and discussing before we could agree that the third *m* is the 7th value of our data set. When the pupils had discovered that it did not matter from which side we did the counting for the 7th value they all agreed that what we had reached was in fact the middle value. Everybody now completed his or her plot by inserting the missing *m*-value.

To round off everyone was supposed to write something that one can see from the stem-and-leaf display and the *mmm*-plot. One typical comment is the following: "I find the rivers to be very long. The shortest is 27 mil and the longest is 72 mil (of the ones we investigated)."

Some teachers at an in-service course criticised the activity as being without interest and forcing the pupils to use a specified way of reporting, not being given the opportunity to decide for themselves how to display the data.

I myself find the activity worthwhile. It has taught us something about the rivers in the our country, and it has made us use and think about numbers in some different ways. Some other time the pupils will be given the opportunity of choosing their own data, of deciding what kind of displays to use. I do not mean to say that stem-and-leaf displays and *mmm*-plots are the only worthwhile ones.

Rolf Biehler, one of the pioneers in Europe among users of EDA in mathematics education (Biehler, 1982), has also used the idea of an intermediate plot before introducing the full box-plot (Biehler, 1990).

Several teachers of in-service courses who have tried box-plots with their 11–12-year-olds have reported that their pupils have had no difficulties at all in grasping the ideas behind them.

One of those teachers, Erland Johansson of Gammelstad, Sweden, reported that his Grade 5 class was not at all impressed by the stem-and-leaf display but got very enthusiastic about the box-plot, which, by the way, Erland introduced successfully without any intermediate *mmm*-plot.

Another teacher, Peter Palo of Kiruna, Sweden, reported that, although he had covered stem-and-leaf displays as well as box-plots with his Grade 8 class, none of the pupils chose these options when left to decide themselves on the kind of diagram to use. In Peter's case the reason might be that his pupils had not spent enough time on stem-and-leaf displays and box-plots, and so had not yet acquired a working knowledge and enough confidence, when they were to select an appropriate method.

But, then, the most important point is not that one uses a particular diagram, as long as the choice is reasonable, what matters most is that the learner is active and cares about the data.

As a follow-up to the 13 longest rivers of Sweden we looked at the 13 longest rivers of the whole world.

The data we had available gave the lengths in km with 4 digits. We discussed the possibility of shortening the numbers somehow, before attempting to make a stem-and-leaf display, and the pupils suggested rounding, which was something they knew a little about.

I proposed reducing the values to two digits by simply cutting, which led to a discussion about estimation and place value, again worthwhile and important.

Note that cutting has its obvious advantages when estimating mentally. Barbara and Robert Reys call this efficient technique *front-end estimation* (see, for example, Reys, 1986). The discussion also gave us reason to talk about orders of magnitude, which I consider one crucial component of a fully developed number sense. In particular when we compared the lengths of the 13 longest Swedish rivers with the 13 longest in the world, the significance of the concept of order of magnitude became clear.

5. A look to the future

"A basic problem about any body of data is to make it more easily and effectively handleable by minds - our minds, her mind, his mind." (John W Tukey, 1977)

During the next school year all Grades 1–3 of Porsöskolan, Luleå, Sweden, which is in the north, will exchange letters with corresponding classes in the south of Sweden. Among other things the children will exchange data, for example, temperature readings. They will send diagrams to each other and they will interpret each other's plots. We will focus on communication. I hope these activities will generate much to report on some future occasion.

What I have reported here are experiences from my teaching. There is much more to be done here in the linking of statistics to the formation of the number concept. What is clearly lacking so far are field studies that, somehow or other, reveal the impact of this approach in suitable developmental terms. The other day I happened to meet Daniel of 5A, the class where the pupils guessed the length of the Lule river. I asked him if he remembered that length, and he frowned and asked if I did not know that the summer holidays had already started. I replied that I did know, but that I was too curious not to ask. He said, "I think it was ... er ... 425 kilometres." He was 25km off the actual value, for me that was a satisfactory evaluation, not very scientific though.

References

Biehler, Rolf (1982) *Explorative Datenanalyse.* Materialien und Studien Band 24, Institut für Didaktik der Mathematik der Universität Bielefeld.

Biehler, Rolf (1990) *Using an intermediate step in the introduction of box-plots.* (Personal communication)

D'Ambrosio, Ubiratan (1986) Socio-cultural bases for mathematical education. In: Marjorie Carss (ed) *Proceedings of the Fifth International Congress on Mathematical Education,* Birkhäuser, Boston, 1-6.

Dunkels, Andrejs (1986) EDA in the primary classroom - graphing and concept formation combined. In: *Proceedings of the Second Conference on Teaching Statistics (ICOTS II),* Victoria BC, 61-66.

Freudenthal, Hans (1973) *Mathematics as an Educational Task.* D Reidel Publishing Company, Dordrech-Holland.

Kamii, Constance (1985) *Young Children Reinvent Arithmetic. Implications of Piaget's Theory.* Teachers College Press, New York.

Kamii, Constance (1990) Constructivism and beginning arithmetic (K-2). In: Thomas J Cooney and Christian R Hirsch (eds) *Teaching and Learning Mathematics in the*

1990s. 1990 Yearbook. National Council of Teachers of Mathematics, Reston, Virginia, 22-30.

Kitcher, Philip (1984) *The Nature of Mathematical Knowledge.* Oxford University Press, New York.

Neuman, Dagmar (1987) The origin of arithmetic skills. A phenomenographic approach. *Göteborg Studies in Educational Sciences* **62**. Acta Universitas Gothoburgensis.

Pereira-Mendoza, Lionel and Dunkels, Andrejs (1989) Stem-and-leaf plots in the primary grades. *Teaching Statistics* **11(2)**, 34-37.

Reys, Barbara (1986) Teaching computational estimation : concepts and stategies. In: Harold L Schoen and Marily J Zweng (eds) *Estimation and Mental Computation. NCTM 1986 Yearbook.* Reston, Virginia, 31-44.

Tukey, John W (1977) *Exploratory Data Analysis.* Addison-Wesley, Reading, MA.