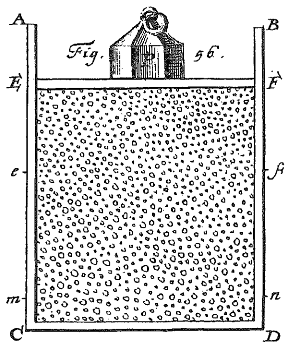
Historical and didactical phenomenology of the average values

BAKKER, Arthur
Utrecht University (Nederland)

Abstract

This study was carried out as a preparation to the development of instruction material for statistics. The history of statistics was studied with special attention to the development of the average values: the arithmetic, geometric, harmonic mean; median, mode and midrange. Also sampling and distribution are discussed. After an introduction on phenomenology, this article firstly discusses a so-called historical and then a didactical phenomenology of the average values.

The average values form a large family of notions that in early times were not yet strictly separated. There are many parallels between history and the development of students' conceptions. It appears to be important that students discover many qualitative aspects of the average values before they learn how to calculate the arithmetic mean and the median. From history, it is concluded that estimation, fair distribution and simple decision theory can be fruitful starting points for a statistical instruction sequence.



Modèle de gaz ayant permis le traitement statistique de la théorie cinétique des gaz, tiré de l'*Hydrodynamica* de Daniel Bernoulli.

1 Introduction

As a preparation to the development of statistical instruction material for 12-year-old students, I studied the early history of statistics. The reason for this is that I assumed that there are parallels between history and education that give inspiration for hypothetical learning trajectories. These learning trajectories are developed with the aim that students somehow reinvent the mathematical concepts (FREUDENTHAL 1973 and 1991). The Dutch mathematician and historian DIJKSTERHUIS (1990) even believed that students recapitulate history at a higher speed. Problems that students encounter when learning mathematics resemble the problems that former generations of mathematicians dealt with. With this in mind, I hoped to find the historical contexts that led to the development of statistical notions, and to detect the conceptual obstacles that mathematicians and users of mathematics encountered. The historical insights are discussed in relation to the results of four exploratory field tests.

1.1 Phenomenology

In FREUDENTHAL's *Didactical Phenomenology of Mathematical Structures* (1983) we find a method of studying the relations between mathematics, history and education. Freudenthal makes a distinction between *phainomena*, phenomena that we want to understand or structure, and *nooumena*, the entities of thought with which we organize these phainomena. Mathematical concepts are examples of such *nooumena* with which we organize our experiential world. This view explains his special way of defining phenomenology:

> Phenomenology of a mathematical concept, a mathematical structure, or a mathematical idea means, in my terminology, describing this *nooumenon* in its relation to the *phainomena* of which it is the means of organising, indicating which phenomena it is created to organise, and to which it can be extended, how it acts upon these phenomena as a means of organising, and with what power over these phenomena it endows us. If in this relation of *nooumenon* and *phainomenon* I stress the didactical element, that is, if I pay attention to how the relation is acquired in a learning-teaching process, I speak of *didactical* phenomenology of this *nooumenon*. (...) if "is ... in a learning-teaching process" is replaced by "was ... in history", it is *historical* phenomenology. (1983, pp. 28-29).

In the context of my article, I use the following simplified definitions:

- *Phenomenology* of a mathematical concept is the study of the relation between that concept and the phenomena it organizes.
- *Historical* phenomenology is the study of the historical contexts in which certain mathematical concepts arose in order to understand why these arose.
- *Didactical* phenomenology is the study of the relation between the mathematical concepts and the phenomena in which they arise with respect to the process of teaching and learning. See also (GRAVEMEIJER 1994, p. 95).

This suggests studying statistics from two perspectives: a historical and a didactical perspective. In the following I will start with a section on the historical phenomenology of the average values. The examples are ordered chronologically. After that, I will discuss basically the same phenomena from a didactical point of view in the same order.

Philosophically seen, there is a difficulty when we make the distinction between *phainomenon* and *nooumenon*: it is not possible to see the phenomenon separated from concepts, since a concept determines and influences the phenomenon. For an educational purpose, this is not a major problem, as long as we keep in mind that students need not see the same phenomena as we do, with our understanding of certain concepts. Studying history can help us to see certain phenomena through the eyes of people who did not have the same concepts. This may help us to understand the learning process of students.

1.2 Historical phenomenology of the average values

The historical study of the average values is difficult for three reasons. First of all, history is often written with respect to men or books, not with respect to concepts. Second, most historical studies start around 1660, when statistics and probability were born according to Kendall (PEARSON & KENDALL 1970, p. 45) and HACKING (1975). Mathematical statistics is even younger, namely from the late 19th century. For my purpose it was necessary to go back to the very start of statistical inference, because my target group is young children. So where I write 'statistical' some readers might prefer to read 'pre-statistical'. Third, statistics is a subject (unlike probability) that was mainly born outside mathematics and was often not seen as part of mathematics. Its birth and growth was due to astronomy, demography, economics, medicine, genetics, biometry, anthropology, the social sciences and many other areas. This might explain why many 'histories of mathematics' pay very little attention to statistics.

1.2.1 Estimation

The oldest example of using an implicit kind of average I have found concerns the estimation of the number of leaves and fruit on two great branches of a spreading tree in ancient India (HACKING 1975, p. 7). How did Rtuparna, the protagonist of the story, do this? He estimated on the basis of one single twig, which he multiplied by the estimated number of twigs on the branches.

I assume he chose a typical or an average twig. I see this as an implicit use of (an intuitive predecessor of) the arithmetic mean, since one number represents all others and this number is somehow in the middle of the others; it is less than the greater, and greater than the smaller numbers; what is too much on the one hand is too little on the other. This use of an average has therefore to do with compensation and balance.

Another example of estimation is a passage written by the Greek historian Herodotus (485-420 BC) on the Egyptians (RUBIN 1968, p. 31):

> They declare that three hundred and forty-one generations separate the first king of Egypt from the last mentioned [Hephaestus] – and that there was a king and a high priest corresponding to each generation. Now reckon three generations as a hundred years, three hundred generations make ten thousand years, and the remaining forty-one generations make 1,340 years more; thus one gets a total of 11,340 years...

The important point in this quotation is the assumption that three generations is a hundred years. This assumption was made to estimate the total amount of years between the first Egyptian King and Hephaestus.

In daily life we often use this kind of average when estimating things. We can think of estimating the number of people coming to a party, the number of bottles we need for that, and the price of our shopping et cetera.

1.2.2 Thucydides: majority, average and midrange

Other very old examples of statistical or pre-statistical reasoning can be found in the work of one of the first scientific historians, Thucydides (460-400 before Christ). The following quotations are from his *History of the Peloponnesian War*. The reader is invited to decide how he or she would translate these two episodes into modern statistical terms (RUBIN 1971, p. 53):

> The problem was for the Athenians ... to force their way over the enemy's wall. Their method was as follows: they constructed ladders to reach the top of the enemy's wall, and they did this by calculating the height of the wall from the number of layers of bricks at a point which was facing in their direction and had not been plastered. The layers were counted by a lot of people at the same time, and though some were likely to get the figure wrong, the majority would get it right, especially as they counted the layers frequently and were not so far away from the wall that they could not see it well enough for their purpose. Thus, guessing what the thickness of a single brick was, they calculated how long their ladders would have to be...

> Homer gives the number of ships as 1,200 and says that the crew of each Boetian ship numbered 120, and the crews of Philoctetes were fifty men for each ship. By this, I imagine, he means to express the maximum and minimum of the various ships' companies ... If, therefore, we reckon the number by taking an average of the biggest and smallest ships...

In the first example, we could see an implicit use of the mode, here indicated by 'the majority'. Note that 'the majority' probably means 'the most frequent value' and not 'more than half'. So in this situation, the Greeks assumed that the most frequent number would be the correct one. In order to find the total height of this number of bricks, they needed another estimation, supposedly of the expected or the average height of a brick. Note that in the first example and the Indian one, it is the total number that counts; they are not primarily interested in the average. The implicit average value is just used to find the height of the wall or the number of leaves.

In the second example, we again see an estimation with the help of an average value. It seems that Thucydides interprets the given numbers as the extreme values, so that the total amount of men on the ships can be estimated. He suggests that this be done by taking the average of these two extremes. In fact this is called the midrange: the midrange is defined as the arithmetic mean of the two extremes. RUBIN[1] (1971, p. 53) writes about this:

> This technique of averaging the extreme values of the range to obtain the arithmetic mean or mid-range can be justified if certain assumptions are defensible, i.e., that the underlying distribution is at least approximately symmetrical or rectangular.

Resuming, we encountered certain phenomena, or problems, that were organized by certain intuitive variants or predecessors of contemporary concepts. When estimating the number of years between two important persons, Herodotus used a number that we now would call the *average*. And Thucydides used a method that we would call *taking the midrange*.

[1]In the American Statistician, RUBIN discussed Adam Smith (1959), Malthus (1960), Karl Marx (1968), Herodotus (1968), Thucydides (1971), Darwin (1972), a medieval household book (1972), and Shakespeare (1973) from a statistical point of view.

1.2.3 Majority, voting and democracy

A very implicit use of statistics is in voting: the majority (the most frequent or common value) is seen as representative of the population. The majority rules. This is the basis of democracy. Also in the Bible and Talmud there are rules like 'follow the majority'. In the Talmud we can read for instance:

> In the entire Law we adopt the rule that a majority [or the larger portion] is equivalent to the whole. (RABINOVITCH 1973, p. 38).

From a modern point of view we might think of the *mode*[2].

1.2.4 The Greek definition of the arithmetic mean

In Pythagoras' time, three kinds of means were known in Greece: the arithmetic, geometric and harmonic mean (HEATH 1921, p. 85; Iamblichus 1939). The theory of these mean values was developed in his school with reference to music theory and arithmetic. Consider for example the musical proportions 6:8:9:12. 8 is the harmonic mean between 6 and 12, and 9 is the arithmetic mean. The proportions 6:8 = 9:12 (a fourth as a musical interval), 6:9 = 8:12 (a fifth), 6:12 (an octave) all form consonant intervals; 8:9 is a second.

Greek mathematics had a different form and aim than modern mathematics. Greek mathematics, even number theory, was highly geometrical and visual. Numbers were represented by lines. This difference between Greek and modern mathematics can be illustrated with the different definitions of the arithmetic mean. The Greek definition was: the middle number b is called the arithmetic mean if and only if $a - b = b - c$. Note that this definition differs from the modern one, $(a + c)/2$, and that it only refers to two values. The Greek version concentrates on the intermediacy, but is difficult to generalize, whereas the modern version highlights the calculation, and is easy to generalize. With education in mind, it is important to note that the Greek definition shows other qualitative aspects than the modern quantitative one. From the Greek definition we can immediately see that the mean is halfway between the two other values. In the didactical section I will come back to this point.

1.2.5. Aristotle's Doctrine of the Mean

Aristotle defines in his Nicomachean Ethics a more philosophical form of the mean, namely the *mean relative to us*. With this notion he explains what virtue is. About the difference between the arithmetic mean and *the mean relative to us* he writes:

> By the mean of a thing I denote a point equally distant from either extreme, which is one and the same for everybody; by the mean relative to us, that amount which is neither too much nor too little, and this is not one and the same for everybody. For example, let 10 be many and 2 few; then one takes the mean with respect to the thing if one takes 6; since 10-6 = 6-2, and this is the mean according to arithmetical proportion [progression]. But we cannot arrive by this method at the mean relative to us. Suppose that 10 lb. of food is a large ration for anybody and 2 lb. a small one: it does not follow that a trainer will prescribe 6 lb., for perhaps even this will be a large portion, or a small one, for the particular athlete who is to receive it; it is a small portion for Milo, but a large one for a man just beginning to go in for athletics.

[2]We find much more about these matters in RABINOVITCH (1973: e.g. section 3.1), a very interesting book on statistical inference in ancient and medieval Jewish literature.

Later in the section he writes about virtue:

> Virtue, therefore, is a mean state in the sense that it is able to hit the mean. (N.E. book II, chapter vi).

From this we may conclude that Aristotle generalized the notion of the mathematical means to situations in daily life. For him, the mean relative to us was an ethical ideal. Another interesting point is that the mean has to do with balance and intuitively means 'not too much and not too little'. This 'definition' is one that students used in all three field tests in which they estimated the number of elephants in a picture. When these students explained their strategies, they defined 'an average box' in a grid as a box in which were 'not too many and not too little' elephants.

1.2.6 Decision theory and Jewish Law

To understand the oldest Jewish examples it is helpful to have some insight into rabbinical Law. First, the Jewish Law is considered a rational pursuit. Although rabbis accepted Divine guidance they insisted on rational methods in coming to decisions. Statistics in this context was solely a decision theory (RABINOVITCH 1973, p. 140). Second, Jewish Law deals mainly with social, ethical and ritual duties. It is not primarily concerned with quarrels and punishing wrongdoing. Rabbis had to decide for example whether food was kosher and how inheritances had to be divided. Gambling was not mentioned since it was considered pagan. I will give one example on kosher food: if 9 out of 10 shops in a city sell kosher meat, and you find a piece of meat in that city, you may consider it kosher (*op. cit.*, p. 45).

A very interesting example concerning multiplicative reasoning and sampling (RABINOVITCH 1973, p. 86) in the Mishnah is found in the Taanit 21a. It concerns whether or not an epidemic has taken place:

> A town bringing forth five hundred foot-soldiers like Kfar Amiqo, and three died there in three consecutive days - it is a plague... A town bringing forth one thousand five hundred foot-soldiers like Kfar Akko, and nine died there in three consecutive days - it is a plague; in one day or in four days - it is not a plague.

In this example, three points are interesting. First, we see that rabbis reasoned proportionally to the total population. Second, a kind of sampling is used: the amount of foot soldiers is used as an indicator of the total population, assuming that foot soldiers form a constant percentage of the population. Third, the rabbis seemed to know normal (or average) death rates: if nine died in one day then it need not be a plague, and also in the case of four days there can be another reason for the nine deaths. So they took into account how the deaths were distributed over the days.

1.2.7 Average as a kind of insurance: fairness and redistribution

In *The World of Mathematics*,MORONEY (1956, p. 1464) writes the following on the history of the average:

> In former times, when hazards of sea voyages were much more serious than they are today, when ships buffeted by storms threw a portion of their cargo overboard, it was recognized that those whose goods were sacrificed had a claim in equity to indemnification at the expense of those whose goods were safely delivered. The value of the lost goods was paid for by agreement between all those whose merchandise had been in the same ship. This sea damage to cargo in transit was known as 'havaria' and the word came naturally to be applied to the compensation money, which each individual was called to pay. From this Latin word derives our modern average. Thus the idea of an average has its roots in primitive insurance.

This view is confirmed by several (etymological) dictionaries (SKEAT 1882) and by WALKER (1931). The Dutch reader may think of 'averij' and the French reader of 'avarie', which is the damage to a ship or to its load after a storm. As a legal term it still refers to the costs of compensating the damage to the load. Here we see that the average has to do with fair redistribution.

1.2.8 Navigation and astronomy: from the midrange to the general form of the arithmetic mean

The *midrange*, the arithmetic mean between the extremes, turns out to be a predecessor to the arithmetic mean (9th - 11th century; Eisenhart 1974, p. 31). Not until the 16th century was it recognized that the arithmetic mean can be extended to n cases: $(a_1 + a_2 + \ldots + a_n)/n$. When sailors in these days had to determine their position on earth, they often used the midrange. Especially when a ship rocks and the compass needle varies, they had to make many observations or look for a while at the compass and find the middle value. Nowadays we know that many observations and errors follow the normal distribution. So the midrange probably was a sensible value to take.

The method of taking the mean for *reducing observation errors* was mainly developed in astronomy (PLACKETT 1958). In astronomy, we want to know a real value, but we can't measure this value directly. We assume that sum of the errors add up to a relatively small number when compared to the total of all measured values.

It is important to note that until then, and long after that, the notion of the mean was qualitative, rather than quantitative. It gradually became a common method to use the arithmetic mean to reduce errors (PLACKETT 1958, EISENHART 1974). It was used to measure, for example, the diameter of the moon, but also used when weighing gold and silver coins.

I would call this insight into the change of definition and the possibility of generalization of the modern arithmetic definition a typical result of historical phenomenology.

1.2.9 Representative value

It took a long time before the mean was used as a representative or substitute. In the examples up till now, the mean has always been used to find some real value: the number of leaves on branches, how many years passed, what the height of the wall was, how many men there were on the ships, and so forth. The only exception seems to be Aristotle's doctrine of the mean, which is a more philosophical variant of the mean.

The Belgian statistician Quetelet (1796-1874) was one of the firsts to use the mean as the representative value for an aspect of a population. Quetelet, as a student of Laplace, was inspired by the method of physics and called his subject 'social physics'. I assume that he could also have been inspired by Aristotle's doctrine of the mean.

In Quetelet's eyes, the average man (*l'homme moyen*) was the ideal man[3]. Cournot (1801-1879) challenged this view; he pointed out that the mean taken for each side from a great number of right triangles could in no way represent the type of a right triangle, since it would almost certainly not be a right triangle at all (PORTER 1986, p. 172). Also Galton and Darwin disagreed with Quetelet, since they were interested in the deviations from the mean. They both were interested in exceptions and they thought that average people were mediocre. Galton was interested in inheritance of genius; in his book *Hereditary Genius* he introduced regression to the mean. By this he meant that the intellectual ability of children of geniuses, but also of 'stupid' people, were generally closer to the mean. For Darwin, deviations from the mean were important since they are a condition for evolution.

Note that the transition from the real value in astronomy to the ideal value of Quetelet, which is a mathematical construct in the social sciences, was an important conceptual change, which is another result of historical phenomenology. The didactical implications will be discussed in section 2.13 and 3.9.

1.2.10 Sampling

We have already seen examples in which sampling was involved: for instance the estimation of the number of leaves and fruit on two great branches (2.1) and the question if there was a plague (2.6). Many other examples of sampling can be found in the Bible and Torah (RABINOVITCH 1973).

Here we only discuss a nice example of simple secular sampling and quality control: the trial of the Pyx in England (STIGLER 1977). At the Royal Mint of Great Britain gold and silver coins were made. Starting from the 12th century, every day one of these coins was put in the Pyx, a box in Westminster Abbey. After a few months or years, the Pyx was opened and the coins were investigated on weight and pureness. If they turned out to be good, this fact was celebrated with a banquet. Otherwise the coinmakers were punished. This is the first example of quality control I have found. For a statistician it is interesting to see that the tolerance interval did not depend on the number of coins; it was a constant percentage of the weight.

Until 1900, samples of the population were considered dishonest and imprecise. Everybody had to contribute to an investigation. When the Norwegian Kiaer presented the representative sample in 1895, he met a lot of resistance. The method was not accepted until 1903, and even in the following years this method of taking representative samples by stratifying was not very successful. The Pole Neyman proposed the random sample in 1934. After that, sampling became increasingly accepted (BETHLEHEM & DE REE 1999).

1.2.11 Median and mode: a few dates

Cournot seems to be the first who used the term *médiane* for the middle value in 1843 (Monjardet in: FELDMAN et al. 1991; PORTER 1986). The English term *median* was coined by Francis Galton (1822-1911) in 1883, who preferred the middle value because it was more robust than the mean, and easier to determine. He also preferred to use the quartiles instead of the standard deviation. Before he coined the median, Galton used the word 'middlemost value' (in 1869) or 'medium' (1980). In a lecture of 1874 he gave the following description but not the current name:

> The object then found to occupy the middle position of the series must possess the quality in such a degree that the number of objects in the series that have more of it is equal to that of those that have less of it. (WALKER 1931, p. 87)

Independently from Galton, Fechner (1801-1887) defined the median but called it *Centralwerth* or C in 1874 (WALKER 1931, pp. 86, 88, 184). WALKER (1931, p. 84):

> In contrast to the practical interest in anthropology which impelled Galton to use the median and related measures, the incentive which Fechner led to discover the median seems to have been a theoretical interest in generalizing the measures of central tendency.

Why was the median called the median? We can find hints for that question in geometry and in the works of Laplace on (continuous) probability density functions: the median is the value such that the left and right parts have equal area. The statistical median of a row of numbers is also in the middle: there are as many numbers left as there are right from the median. If we think of geometry we can imagine another reason why it is called the median: in a triangle, the median is the line that goes to the *middle* of the opposite side and the two parts left and right from the median have equal area. The mode was coined by Karl Pearson in 1894. Fechner spoke of *der dichteste Werth* (D) in 1878. The difficulty with these historical dates is that the notions were used long before in an informal sense, but without being given a name. See for instance the sections on Thucydides and democracy, and the quotation of Galton in this section.

1.2.12 Box plot and stem-and-leaf diagrams

John TUKEY is a famous name connected to exploratory data analysis. He is the inventor and propagator of the box-and-whiskers plot and stem-and-leaf diagram, which became well known in the 1970s (TUKEY 1977). Box plots are suitable for comparing uni-modal distributions. Stem-and-leaf-diagrams can be useful when we have to put numbers into order by hand and if we want to see how they are distributed.

1.2.13 Cultural knowledge

The box plots and stem-and-leaf diagrams are examples of mathematical objects that came late in history, but can come early in education. So we can not simply follow the path of history. This would also be a waste of time, since a lot of mathematical knowledge nowadays is also cultural knowledge. Many children already know in some intuitive sense what *average* and *random* mean and what a survey is, just because they meet these words in their environment. From history, we might conclude that random sampling is a difficult concept because it was only accepted in the 20th century, but this does not imply that students nowadays have the same problems with that notion as people around 1900. Therefore, when making instruction material with history in mind, we should not forget that we could benefit from the students' cultural knowledge.

1.3 Didactical phenomenology of the average values

The phenomena and concepts that were discussed in the historical context will now be discussed in relation to the process of teaching and learning. The order will be basically the same. The

[3] This is a simplified view; in fact Quetelet's view was much more subtle.

protocols come from an exploratory field test with 26 students of age 12 in April 1999. I tested my instruction material in three classes (first year in secondary school) in the period from September to November 1999.

1.3.1 Estimation

The earliest historical examples had to do with estimation. This gives rise to the following question: is it a good idea to start the teaching of the mean with estimations? More than the calculation of the mean can estimation direct the attention to more qualitative aspects of the average: representative, typical, normal, somewhere in the middle, balance, compensation. Generally speaking, estimation does not direct the attention to the average but to the total. I chose to start my instruction material with a picture of a herd of elephants and ask: how many elephants are there in the picture? The 12-year-old children used four main strategies with many variants:

1. Make groups, guess how many are in each group and add, or
2. make a group with a fixed number and estimate how many groups fit into the whole, or
3. make a grid, choose an 'average box' and multiply by the number of boxes in the grid, or
4. count the number of elephants in the length and width and multiply these (Mr. Bean's method of counting sheep).

Strategy 3, which was used most often, is indeed based on an intuitive idea of the average. Discussing why this strategy works students tend to define an 'average box' as one that does not have too many and not too little in it. Compare this with Aristotle's characterization of the mean (2.5).

1.3.2 Translation of terms: majority

Sometimes, the majority is right; see Thucydides on counting the number of bricks (2.2). But in the case of estimating the elephants nobody was right: all students, except the few who used strategy 4, estimated too low. Does the majority have to do with the mode? Indeed, I would anachronistically translate 'the majority' into the mode or modal class. The mode is not a very robust[4] or useful value in statistics. Therefore, I chose to concentrate in my instruction material on the mean and median.

As we saw from the quotations of Thucydides (2.2), it is rather difficult to make implicit aspects of average values explicit. Did Thucydides really think of the midrange in the second quotation? In my field test I encountered a similar difficulty of translating the arguments of students into statistical terms. Often, it was hard to detect the underlying principles of the students' arguments. From only 26 students I got a rich variety of answers to the question what the average is:

[4] The mode is not robust since it is sensitive to outliers. Neither is the midrange. The median though is very robust, even more than the arithmetic mean.

Jennifer:	Yes, the half. The whole, and in between the half, that is the mean.
Charissa:	Everything together.
Bart:	You look between the highest and lowest.
Centina:	The most.
Claire:	What you think it is roughly.
Lisa:	The mean is about a bit in balance.
Kerster:	In between.
Frank:	The midpoint.

Many students said: add and divide by the number.

Keeping the same order, we could translate these answers into: one part of the algorithm (divide by 2), the other part of the algorithm (add everything), midrange, mode, estimation, balance, intermediacy, median or center of gravity and the complete algorithm.

If we turn to philosophy of language, we could see the average values as a large family of related concepts of which the arithmetic mean is just one member (WITTGENSTEIN 1984). Early in history, but also in the development of a child, we see no clear distinctions between all aspects of these concepts. When people organize their world and solve problems, they are urged to become clearer and define more precisely. For the instruction material, we must therefore choose contexts that really ask for clear distinctions. For example, a skewed distribution can show the limitation of the midrange and ask for another measure of central tendency.

1.3.3 The Greek definition: intermediacy

What I conclude from the transition from the visual 'intermediacy' definition of the Greeks (b is the mean of a and c if and only if $a - b = b - c$) to the general form (16th century) is that we should start with a very visual and qualitative notion of average values before teaching the calculation of the arithmetic mean. The students should at least know by experience that the mean is between the extremes. I have noticed that many students forget this aspect if they calculate the mean and sometimes get an answer out of the range. An illustration of this:

Arthur:	How would you estimate the average annual temperature in the Netherlands from this graph or table?
Jennifer:	Add everything.
A:	And then?
Lisa:	Divide by 2.
A:	Why divide by 2?
L:	Because that is the average.

They got 55 degrees Celsius and did not realize that this was very hot.

If students have developed some visual intuition, this might happen less.

The first computer minitool[5] I use for instruction in exploratory data analysis (www.fi.uu.nl/arthur Minitool 1) is designed so that magnitudes are represented by bars. It

[5] The minitools were designed for a teaching experiment of Koeno Gravemeijer, Paul Cobb and others at the Vanderbilt University in Nashville, USA.

turned out that this representation created an opportunity for the students to find the mean by cutting off the longer ones and give these pieces to the shorter ones. When asked to estimate the annual year temperature in the Netherlands from a bar graph, many students spontaneously came up with a compensation strategy. Some said: 'I give a bit of July to January, from August to February and so on'. This visual way of dealing with the mean is also an example of fair redistribution.

1.3.4 Redistribution and fairness

In history, statistical reasoning often had to do with fair division of inheritance or possessions (2.6). Fairness is very important issue to children, so we can profit from that. It can easily be combined with other aspects of the mean, such as compensation and balance. Note that distribution in the sense of division (giving everybody the same amount) is different from distribution in the technical sense (normal distribution, for example).

This is done for example in *Mathematics in Context*, instruction material for American middle schools. Students are asked to rearrange mice in nests in order to answer the question: how many mice does a mother mouse roughly get? If we want to avoid half mice, we could take biscuits instead. The question will then be: how many biscuits does every student get if they are fairly distributed? (See also: MOKROS & RUSSELL 1995; STRAUSS & BICHLER 1988).

1.3.5 Three components

If we consider calculations with the mean, we see three components: the number n, the sum or total Σ and the mean μ. These components can have different roles. When making and analyzing exercises it is useful to categorize the possibilities:

1. Estimation often has to do with finding the total: $n^*\mu = \Sigma$. The used average value mostly stays implicit since the focus is on the total. In this way students get feeling with many aspects of the mean without using it explicitly.

2. Fair redistribution (how much does everyone get?) has to do with finding the mean: $\Sigma/n = \mu$. Sometimes, for example if we want to compare fairly, we need to compensate for the number. By *mean as a measure* is meant the use of the mean as a way of compensating for the number n, e.g. if we use parts per million, a percentage, gross national product per head, et cetera. In other words: the mean is used to make fair comparison possible. In history we see this when a criterion is given whether an epidemic has taken place (2.6).

3. I also asked how many 12-year-old students could go into the basket of a hot air balloon. They got the allowed weight, they estimated the mean weight of students and calculated the number $n : \Sigma/\mu = n$. The same exercise could be done with an elevator instead of a balloon. This exercise also implicitly asks for an average, namely the estimated weight of these students.

1.3.6 Reduction of error

In the counting example of Thucydides the mode is the best value. In astronomy, it often is the mean that is the best value, namely as an approximation of the assumed real value. We only

have data that are approximations of this value, so the mean substitutes the real value. This use of the mean, I fear, is too difficult for students that don't yet have learnt physics at school. We could, however, look for other contexts in which students have some intuition about reducing errors.

1.3.7 From the midrange to the general form of the arithmetic mean

In history, the midrange turned out to be a predecessor of the mean (see 2.8). This suggests allowing the midrange as an average value, for instance in estimation strategies. When estimating the number of elephants in a picture, a few students indeed used the midrange. They counted the box with the least and with the most elephants, calculated the mean of these two, and multiplied by the number of boxes. Of course, the students should experience that the midrange is only useful if the distribution is symmetric, so the following lesson I showed both a symmetric and a skewed distribution to make clear when this method is useful and when it goes wrong. Another disadvantage of the midrange that students should encounter is that it is very sensible to outliers, as can be shown easily.

1.3.8 Balance and midpoint

A very nice way of presenting the mean is by means of a balance. HARDIMAN, WELL & POLLATSEK (1984) have written about the usefulness of the balance model in understanding the mean. For younger students, I fear though, this method is less suitable since they need more knowledge of physics to understand the balance model. Therefore, I preferred to start in my field tests with estimation and redistribution.

1.3.9 Representative value

The aspect of representativeness comes very late in history. It is rather a huge step from the real value, as in astronomy, to an ideal or representative value, as in the social sciences. Quetelet, the man of the *l'homme moyen*, is an example of the latter use. We can't simply say that this latter use is more difficult than the former one. There are several layers of understanding the mean as a representative value. Students have no problem in seeing an average Dutchman as a normal Dutchman, but have more difficulties with artificial constructs like the average size of a family, which is a decimal number.

1.3.10 Sampling

There are several layers in the understanding of sampling. The historical examples can help to define these layers. For example, the Trial of the Pyx is a very elementary form of sampling, whereas random sampling is very abstract. It is clear that sampling should be learned, because students can only understand certain data if they know how these data were created.

Note that sampling in the form of surveys is a common practice in society, so many children intuitively know what random means and what a sample is. Therefore, didactics need not be parallel to history (2.13).

1.3.11 Median and mode

In most statistics courses distributions are described in relation to the mean and standard deviation. For younger students, I prefer the median and quartiles, because they don't know the square root yet. Later on, of course, they should learn the mean and standard deviation, since these are used more often and have great theoretical advantages.

The students should also meet situations where using the median or mode makes more sense (e.g. salaries).

1.4 Conclusions

Historical phenomenology can be helpful when designing instruction material. The obstacles in history tell us where the difficult conceptual transitions are and what might facilitate their teaching and learning. Of course, it is not necessary that students learn all the detours that history made. A shorter route is possible because of the cultural knowledge of the students and the experience of teachers with education.

Still, the parallels between history and education are sometimes remarkable. For example, I have detected many aspects of the mean with only one field test with 26 students. Their answers to the question 'what is the average?' already yielded many associations with: parts of the algorithm (adding or dividing), median, mode, balance, midrange, representativeness, compensation, midpoint and intermediacy. All these aspects and different average values can be seen as a large family of concepts that initially have no precise borders or definitions (WITTGENSTEIN 1984). This implies that we should start very broadly and not stick to the arithmetic mean as the only possible average value. In the 19th century, statistics seemed to be the science of arithmetic means (FELDMAN et al. 1991), but many people such as Galton and Darwin have opposed that tendency, and preferred other values such as the median. About the midrange: this concept is useful for the educator and teacher to describe what students do. We could allow students to use it as long as they learn the limitations as well.

When students see a lot of different situations and aspects of the average values, they develop a good basis and motivation to make distinctions between them. Not until then will they appreciate and understand the precise mathematical definitions of the statistical concepts. My most important point is that students should first discover many qualitative aspects of the mean before they learn the general form or the algorithm. I think that estimation and redistribution can be good starting points, especially when the visual aspects are stressed.

Literature

ARISTOTLE. Nicomachean Ethics. Loeb Classical Library (1926/1994), London.

BETHLEHEM, J.G. & S.J.M. DE REE (1999). 100 jaar CBS: van populatie naar steekproef. CBS research paper 9901.

BIEHLER, R. & H. STEINBRING (1991). Entdeckende Statistik, Stengel-und-Blätter, Boxplots: Konzepte, Begründungen und Erfahrungen eines Unterrichtsversuches. Mathematik Unterricht 37(6), 5-32. [There is also an English version]

BOYER, C.B. (1968). A history of mathematics. Revised edition (1991) by U.C. Merzbach. John Wiley & Sons, Inc.

COBB, P. (1999). Individual and collective mathematical development: the case of statistical

data analysis. Mathematical thinking and learning 1(1), 5-43.

DAVID, F.N. (1962). Games, Gods and Gambling. Charles Griffin, London.

DIJKSTERHUIS, E.J. (1990). Clio's Stiefkind. Berkel, K. van (red.). Bert Bakker, Amsterdam.

EISENHART, C. (1974). The development of the concept of the best mean of a set of measurements from antiquity to the present day. 1971 A.S.A. Presidential Address (Unpublished).

FELDMAN, J., G. LAGNEAU & B. MATALON (Eds.) (1991). Moyenne, milieu, centre: histoire et usages. Paris Ecole des Hautes études en Sciences Sociales.

FREEDMAN, D. et al. (1978). Statistics. Norton Company, New York.

FREUDENTHAL, H. (1932). Qualität und Quantität in der Mathematik. Euclides 8, 89-98.

FREUDENTHAL, H. (1966). De eerste ontmoeting tussen wiskunde en de sociale wetenschappen. Brussel.

FREUDENTHAL, H. (1973). Mathematics as an educational task. Reidel, Dordrecht.

FREUDENTHAL, H. (1983). Didactical phenomenology of mathematical structures. Reidel, Dordrecht.

FREUDENTHAL, H. (1991). Revisiting mathematics education. Dordrecht: Kluwer Academic Publishers.

GRAVEMEIJER, K.P.E. (1994). Developing realistic mathematics education, Freudenthal Instituut. Cdβ reeks.

HACKING, I. (1975). The emergence of probability. A philosophical study of early ideas about probability, induction and statistical inference. Cambridge University Press.

HARDIMAN P.T., A.D. WELL & A. POLLATSEK (1984). Usefulness of a balance model in understanding the mean. Journal of Educational psychology, 76(5), 792-801.

HEATH, T. (1921/1981). A history of Greek mathematics. Dover, New York. Iamblichus (1939/1991). In: Greek Mathematics, volume 1. Loeb Classical Library, Harvard University Press, Cambridge, MA.

KENDALL, M.G. & R.L. PLACKETT (1970). Studies in the History of Statistics and Probability, volume II. Griffin London.

Mathematics in Context (1998). Picturing Numbers.Britanica, Chicago

MOKROS, J. & RUSSELL, S.J. (1995). Childrens' concepts of average and representativeness. Journal for Research in Mathematics Education 26(1), 20-39.

MORONEY, M.J. On the average and scatter. See Newman, the World of Mathematics.

NEWMAN, J.R. (1956/1988). The world of mathematics. Volume 3. Tempus Books, Washington.

NICOMACHUS. See Iamblichus, Greek Mathematics.

PEARSON, E.S. & M.G. KENDALL (1970). Studies in the History of statistics and probability. Volume I. Griffin, London.

PLACKETT, R.L. (1958). The principle of the arithmetic mean. In Pearson & Kendall (1970)

PORTER, T.M. (1986). The rise of statistical thinking, 1820-1900. Princeton University Press.

RABINOVITCH, N.L. (1973). Probability and statistical inference in ancient and medieval Jewish literature. University of Toronto Press.

RUBIN, E. (1968). The statistical world of Herodotus. The American Statistician, February 1968.

RUBIN, E. (1971). Quantitative Commentary in Thucydides. The American Statistician, December 1971.

SCHNEIDER, I. (1991). Essay review of books of Daston (1988), Porter (1986) and Stigler (1986). Historica Mathematica 18, 67-74.

SKEAT, W.W. (1882/1972). A concise Etymological dictionary of the English language. Oxford University Press.

SINGER, J. & J.B. WILLETT (1990). Improving the teaching of applied statistics: putting the data back into data analysis. The American Statistician 44(3), 223.
STAMHUIS, I.H. & A. DE KNECHT-VAN EEKELEN (1992). De met cijfers bedekte negentiende eeuw. Erasmus Publishing, Rotterdam.
STIGLER, S.M. (1973). Simon Newcomb, Percy Daniell, and the history of Robust Estimation 1885-1920. Journal of the American Statistical Association.
STIGLER, S.M. (1977). Eight centuries of sampling inspection: the trial of the Pyx. Journal of the Americal Statistical Association, 72, 493-500.
STIGLER, S.M. (1986) . The History of Statistics. The measurement of uncertainty before 1900. Harvard University Press.
STRAUSS, S. & E. BICHLER (1988). The development of children's concepts of the arithmetic mean. Journal for Research in Mathematics Education 19(1), 64-80.
TUKEY, J.W. (1977). Exploratory Data Analysis. Addison-Wesley Publishing Company, Massachusetts.
TUKEY, J.W. (1980). Why we need both exploratory and confirmatory. The American Statistician, 34(1), 23-25.
WALKER, H.M. (1931). Studies in the history of stastistical method; with special reference to certain educational problems. Baltimore, Williams and Wilkins Company.
WITTGENSTEIN, L. (1984). Philosophische Untersuchungen. Suhrkamp.

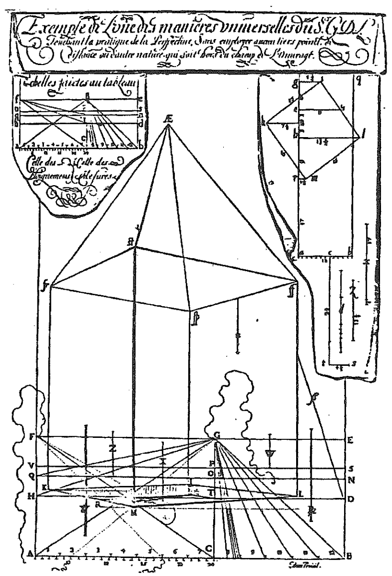


Planche du traité de perspective de Desargues de 1636.

De Brunelleschi à Desargues ou des problèmes liés à la représentation plane d'objets de l'espace ...

BALLIEU, Michel
C.R.E.M. (Belgique)

Abstract

À partir de textes originaux – Leon Battista Alberti, Piero della Francesca, Leonardo da VInci, Albrecht Dürer, Antonio Manetti, Giorgio Vasari, Simon Stevin, ... – nous tentons de mettre en évidence les difficultés qui surgissent lorsque l'on essaie de représenter dans un plan des objets de l'espace. C'est un problème auquel nos jeunes élèves sont confrontés dans leur cours de géométrie de l'espace. Nous tentons de montrer que ce problème mathématique qui va déboucher sur une nouvelle géométrie -la géométrie prospective- trouve en fait son origine dans un idéal profondément humain de recherche d'esthétique en peinture.

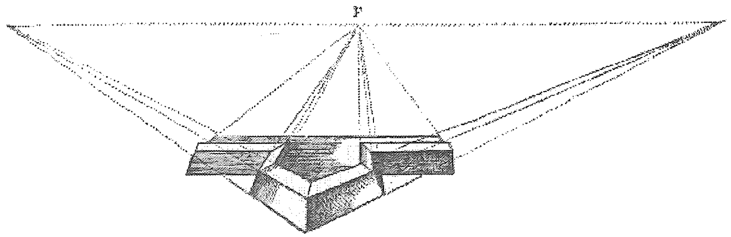


Illustration tirée du *De sterkte Bouwing* de Stevin