Gustav Adolf Lörcher, Freiburg

The Development of Number Sense

Abstract:
In order to find out how number sense develops with increasing age, an estimation test, including numbers of quantities, lengths and weights in different orders of magnitude was developed and given to children in primary schools, in secondary schools and to mathematics students. The results show strengths in estimating lengths (in cm and m) and weaknesses in estimating weights and lengths or quantities of large orders of magnitude, and give an idea of how estimation abilities develop. Some gender differences seem to exist. From the numbers actively used for estimation, some conclusions are drawn about the presumably multiplicative structure of the underlying number concepts.

Introduction

In daily life you have to make many quantitative decisions. The competence you need is called number sense (Dehaene), numeracy (Glenn), or quantitative literacy (Zacharias). How do you acquire this competence? What does mathematics education contribute to the development of number sense or in what way does it fail to contribute to it?

In order to answer these questions it may be useful to look at how children, young persons and adults estimate. In the following I will suggest why number sense is important and which kind of number sense is required. I will consider the role estimation plays in developing number sense and also report some results which young people obtained in the same estimation test given to different age groups in Baden-Württemberg.

The importance of number sense

In all areas of daily life you have to register, to compare, to create, to process, to imagine and to control numerical relations between quantities. For example, avoiding financial disadvantages, finding your way in traffic, optimising your journeys, minimizing risks, making best use of different possibilities, planning your life in the short run and in the long run and judging what may be important in society today or in the future. In all these questions, numbers play a crucial role and impressive examples may be found in the articles of Winter or Paulos.

In order to cope with the abundance of numerical data, you need a personal numerical control system, which enables you to detect and avoid severe inconsistencies, to see if different magnitudes fit together and to judge whether the data given are reasonable. This control system can be developed and improved in maths education by creating a feeling for numbers and by promoting imagination for magnitudes in different areas and of different orders. When, after many years of mathematics education, educated people still confuse millions and billions (cf. Hofstadter), you see that this task is far from being completed.

The development of number sense

In the literature you find a lot about how number sense begins to develop in young children before or after entering school (e.g. Fuson, Gerster & Schultz). However not much is known about how this development continues as children get older.

In the syllabus e.g. in Baden-Württemberg, you will find proposals on how children may acquire a step-by-step understanding of bigger numbers. In the first two grades numbers up to 100 are introduced, in the 3rd grade up to 1000, in the 4th grade up to 1 million and in the 5th grade up to billions and trillions, as well as smaller numbers (in the 6th grade fractions and decimal numbers). The leading principle for developing the number system in the syllabus is not addition (represented in the number line) but multiplication (and division) by powers of 10 (represented in the decimal notation).

Getting familiar with numbers can be compared to the acquisition of the vocabulary of a language. Most new words you meet go into your passive personal vocabulary. Only to a lesser extent are you able to make active use of words, which belong to your considerably smaller active personal vocabulary.

In the same way, the syllabus informs about the passive use of numbers, but gives no hint of how far children and young persons are able to use them actively in different areas and orders of magnitude. The interesting question is how this active number vocabulary develops in association with the kind of measurement and the order of magnitude. Is the personal number vocabulary and its grammar like a number line, where addition and subtraction are the main number operations, or is it organized as a place-value system, where multiplication and division (by 10) are the most important number operations?

During the last few years new light was thrown on these questions by the research of cognitive neuroscientists like DEHAENE and BUTTERWORTH, who have investigated the processing of numerical quantities in the human brain. DEHAENE sees an innate quantitative imagination in new-born human beings and in higher animals that enables them to recognise and to distinguish quantities up to 5 and makes possible rough numerical comparisons of different quantities. These comparisons are not additive in principle but multiplicative and follow Weber's law. Human beings are further able to construct a symbolic representation of quantities.

The multiplicative structure of the recognition of quantities is demonstrated by the place-value notation of numbers, in which the magnitude of a number is expressed by the number of places roughly corresponding to the base ten logarithm of this number.

The importance of estimation

It is still not very well known (cf. BENTON, SOWDER), by which methods this numerical imagination can be trained (e.g. by rounding numbers, calculating with powers of ten or by estimating). In classroom practice however it has turned out that estimating real quantities is both a highly motivating device and a successful way to find out interesting things about the numerical capabilities of people in any age group. In the classroom, estimation questions can be given not only to the pupils but also to the teacher – and it usually turns out that some pupils give better estimates than their mathematics teacher.

In the following, I report about the development of an estimation test and about the main results.

The test was developed within a statistics course at the Pädagogische Hochschule Freiburg. From 1987 to 1997 about 600 mathematics students were given an estimation test. The test was put into its final form in 1998 and since then has been given to about 200 mathematics students, to about 60 children in primary schools and nearly 800 pupils in secondary schools (in all grades of both Hauptschule and Realschule, but not in Gymnasium). In its final form the test included 31 estimation questions (3 about the number of quantities, 20 about lengths and 8 about weights) in different orders of magnitude. Some quantities to be estimated were visual (like the quantity of numerous points on a sheet of paper or the length and breadth of a sheet of paper). Others were easy to imagine (like the diameter of a given coin or the length and breadth of a banknote) and some were only accessible with difficulty, because they (or some numerical impression of them) had to be memorized (like the world population or the distance from the earth to the sun). In selecting and formulating the items it was felt important that they should be accessible to all age groups from primary school to adults. That means that all the objects to be estimated should be well known in daily life.

In its final form, all questions were on one side of a sheet of paper and could be answered without time limitation. Usually it took between 15 and 30 minutes.

A measure of the success rate in estimation

In most existing tests, estimation skill can only be assessed roughly in a qualitative way (cf. MEIßNER). In order to compare estimation skill quantitatively, a new measure for the success rate of an estimation task had to be developed.

The starting point came from the observation that in comparing two quantities most people rather think "$\frac{1}{4}$ more or $\frac{1}{4}$ less" than "10 more or 10 less" and in this way divide the estimated numerical value of the two quantities instead of subtracting. Therefore, for most people, when the estimated value is less than the true value of a quantity, it is reasonable to divide the estimated by the true value and give the percentage.

When the estimated value is bigger than the true value, it seems reasonable to do the same, but trying to compare over- and underestimation leads to severe inconsistencies. The more the estimation exceeds the true value, the greater the inconsistency. For example, if the estimated value is 100% more than the true value, it is twice the true value (in many cases an acceptable guess) whereas an estimated value of 100% less than the true value is 0 and therefore a totally useless result. Overestimating the true value by doubling it seemed to me comparable to underestimating the true value by halving it. So in measuring overestimation you should divide the true value by the estimated value. That means as *a general rule* – always *divide the smaller by the bigger number*.

The result gives the success rate as a proportion, either as the underestimate over the true value or as the true value over the overestimate. In this way you get an estimation quality of 75% not only when the underestimated value is 25% less than the true value but also when the true value is 25% less than the overestimated value (that means an overestimation by $\frac{1}{3}$). Similarly you get a success rate of 25% when the underestimation is $\frac{1}{4}$ of the true value or the overestimation is 4 times the true value.

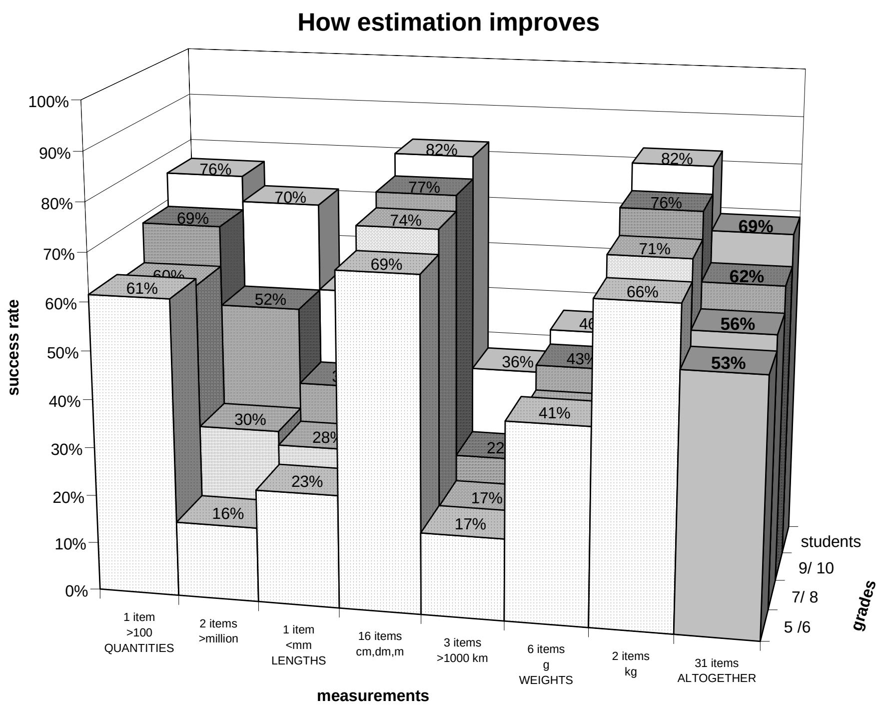
Test results in general

The final form of the test was given to 3 classes in primary school (grade 3 and 4, n = 66), to 35 classes in secondary schools (22 classes in Hauptschule, n = 474; 13 classes in Realschule, n = 309) and to mathematics students in 3 statistics courses (n = 202).

Though estimation is usually neglected in mathematics education throughout all school years, the overall success rate improved during the time spent in school. Whereas pupils in grade 3 and 4 and in grade 5 and 6 attained only about 50%, the success rate increases to more than 60% at the end of secondary school (grade 9 and 10) and rose to nearly 70% with the mathematics students.

age group	grade 3/4	grade 5/6	grade 7/8	grade 9/10	students
n=	66	273	313	197	202
success rate	**.51**	**.53**	**.56**	**.62**	**.69**
standard deviation	.10	.09	.09	.08	.07

The diagram shows the development of estimation success in and after secondary school for different measurements and different orders of magnitude.



Estimation success rates		true	Primary school n=66			Secondary school n=783			Mathematics students n=202		
item	unity	value	mean v.	st.dev.	t-test	mean v.	st.dev.	t-test	mean v.	st.dev.	t-test
number of points		129				.63	.28		.76	.19	
Germany: population	mill.	82	.14	.23		.31	.40	**	.75	.35	**
world population	bill.	6	.16	.28	**	.30	.37	**	.66	.38	**
A4: thickness	mm	.1	.19	.22		.28	.33		.52	.36	
draw a number on the n.l.	cm	.9	.37	.34	**	.58	.32	**	.84	.16	
diameter: 1 Pf	cm	1.7	.53	.21		.55	.22	+	.56	.18	
diameter: 50 Pf	cm	2	.55	.27		.60	.25	++	.64	.19	
diameter: 10 Pf	cm	2.2	.62	.27		.64	.23	+	.68	.18	
diameter: 1 DM	cm	2.4	.61	.27		.69	.23		.76	.16	
diameter: 5 DM	cm	2.9	.60	.29		.72	.24		.80	.16	
drawing a line	cm	5	.55	.38		.75	.31		.86	.18	
10 DM: breadth	cm	6.5	.66	.24		.67	.27		.81	.19	
length of a line	cm	11.9	.78	.24	**	.83	.16	**	.89	.10	
10 DM: length	cm	13	.76	.19		.76	.18		.85	.13	
pos.of a number on the n.l.	cm	16	.57	.39		.75	.33	*	.94	.13	
A4: breadth	cm	21	.73	.31		.75	.27		.89	.12	
A4: length	cm	30	.85	.20		.87	.18		.93	.10	
baby: height	cm	52	.73	.24		.80	.23	++	.89	.15	+
10 years old: height	cm	144	.88	.24		.87	.20		.89	.10	
blackboard: length	m	4	.86	.27		.86	.22	*	.91	.15	
earth: perimeter	Mkm	.04	.14	.21	**	.25	.33	**	.65	.38	**
earth-moon: distance	Mkm	.38	.19	.30	**	.22	.29	**	.28	.31	**
earth-sun: distance	Mkm	150	.04	.12	**	.08	.17	**	.16	.31	**
weight: 1 Pf	g	2	.50	.28		.47	.31	*	.52	.29	
weight: 50 Pf	g	3.5	.43	.31		.41	.28		.48	.26	
weight: 10 Pf	g	4	.48	.25		.44	.29		.51	.27	
A4: weight	g	5	.29	.26	**	.30	.30	**	.32	.30	*
weight: 1 DM	g	5.5	.43	.33		.38	.28		.46	.26	
weight: 5 DM	g	10	.44	.38		.39	.30		.47	.29	
baby: weight	kg	3.3	.48	.32	*	.61	.31		.81	.22	
10 years old: weight	kg	31	.72	.34		.80	.23		.83	.18	
success rate altogether			.51	.10	**	.57	.09	**	.69	.07	**

The most important results are:

Estimation of **quantities** (how many?) improves considerably after grade 8. Before that pupils lack the references to enable them to estimate, for example, the population of a country.

Pupils in grade 5/6 are already able to estimate **lengths** quite well in daily life orders of magnitude (cm, m). The success rate increases during secondary school from about 70% to nearly 80%. For extremely small or large orders of magnitude (thickness of paper, astronomical distances), they do not attain a feeling for the appropriate length during school, and even mathematics students cannot cope successfully with astronomical distances.

As to estimating **weight**s of daily objects weighing only a few grammes, there is practically no gain in success rates during school and after school with the success rate staying below 50%. Body weights (of babies and 10 years olds) are well known at the beginning of secondary school and estimated even better with increasing age.

Special test results

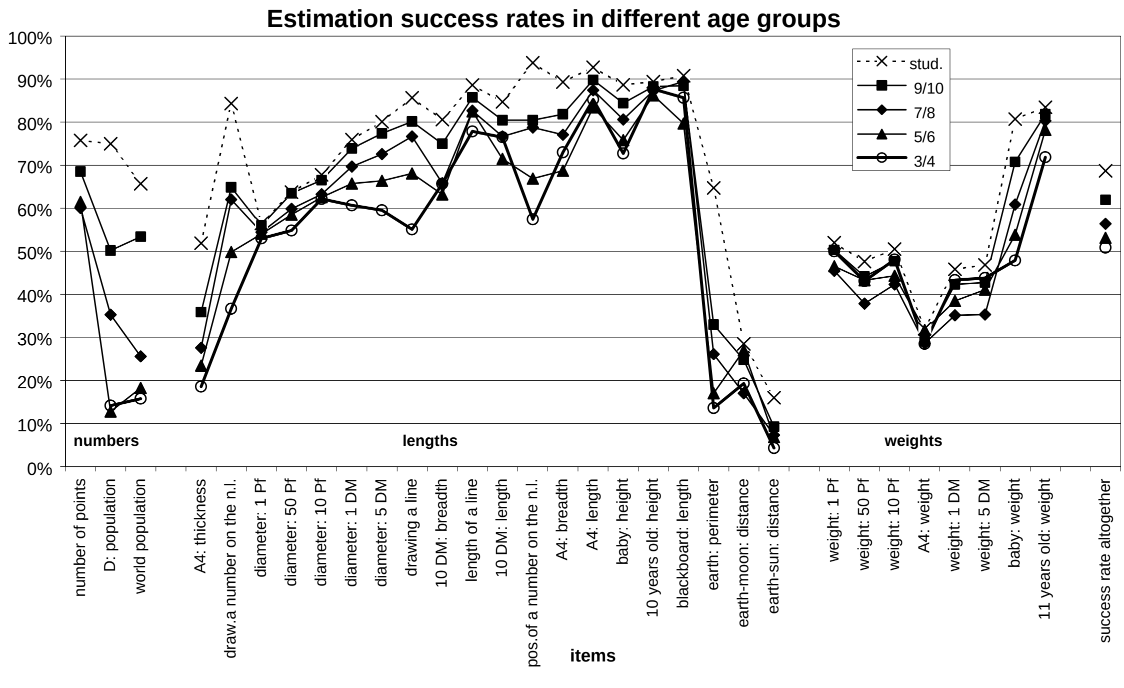
In the table the success rates of the single items are given. The items are ordered according to the type of measurement and the magnitude.

In the t-test column, ** means high significance (p<.01), * means significance (p<.05) in favour of the males and ++ or + in favour of the females.

In most test items there were no significant **gender differences** but in all questions that included big numbers, males estimated significantly better than females. In estimating the position of the number 100 on a number line between 0 and 1000, boys improved their results from 35% in grade 3/4 to 63% in grade 5/6, whereas girls stayed at 39% in grade 5/6 and only improved 2 years later to 59%. Females outperformed boys in estimating the height of a baby and tended in most school years to have better test results when estimating the diameter of coins. In all age groups (with exception of grade 7/8) the overall success rate of males was significantly better than that of females.

As can be seen from the following diagram, the **best results during school time** are attained in estimating the length of an A4-sheet and of a blackboard, both visible during the test, and the height of a 10 year old, a measurement which is easy to bring to mind (all between 80% and 90%). The **biggest gains in school time** are seen to be in estimating population numbers, the position of the number 100 on a number line between 0 and 1000 and, after school, in estimating the thickness of paper and the perimeter of the earth.

The **worst results during school time** occur in estimating small numbers (thickness of paper) or estimating big numbers (especially astronomical lengths). **The lowest gains** both in and after school occur in estimating the diameter of the smallest coin (1 Pf), the distance from the earth to the moon and to the sun and in all estimates of small weights, especially the weight of a sheet of paper.



Over- and underestimation

Measuring the success rate does not give information about whether people tend to estimate too high or too low. If, for each person, the number of underestimated items is subtracted from the number of overestimated items then a positive result indicates overestimation and a negative result underestimation. As the following table shows, a remarkable result arises: primary pupils tend to underestimate. In secondary school this changes for 3 or 4 years but by the time they leave school the majority underestimate again. Nearly ¾ of the mathematics students underestimate.

percentage of underestimators	grade				students	
	3/4	5/6	7/8	9/10	1998-2000	1987-1997
n=	66	273	313	197	202	656
underestimators	61%	50%	40%	64%	73%	75%

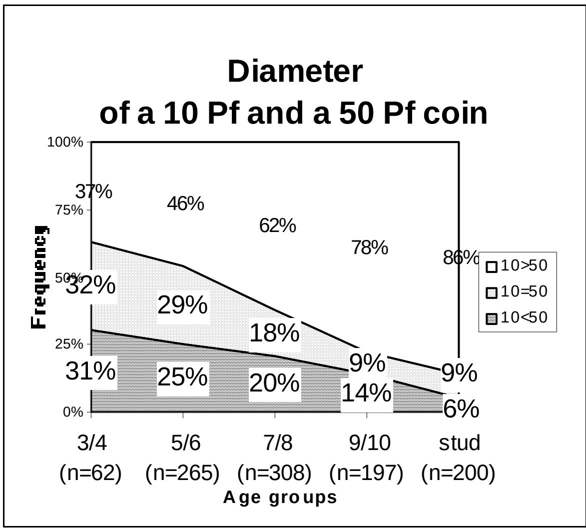
Consistency of estimations

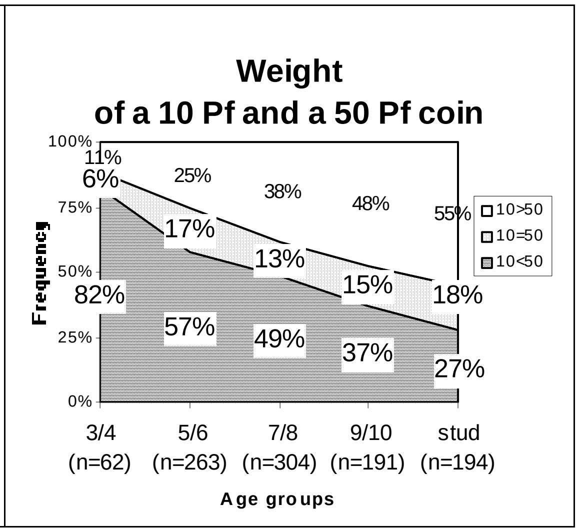
In estimating objects it may happen that people always overestimate or always underestimate, but still have a consistent reference system in which the relations between the objects are adequately represented. So in daily life, for example, it could easily be seen that the length of the 10 DM banknote is twice the breadth, when the banknote is folded in a wallet. In fact, this daily experience does not lead to the expected insight, as the following table shows.

estimation of the format of the 10 DM banknote	grade				students	
	3/4	5/6	7/8	9/10	1998-2000	1987-1997
n=	63	258	291	191	196	625
l < 2b	40%	28%	25%	25%	32%	36%
l = 2b	**32%**	**27%**	**27%**	**34%**	**26%**	**22%**
l > 2b	29%	46%	48%	42%	42%	42%

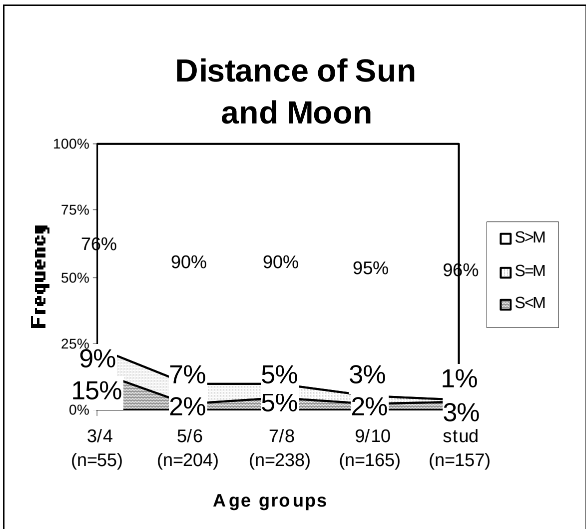
Throughout all the age groups only about ⅓ of those tested realized the double square format of 10 DM. The other ⅔ split about equally into underestimating or overestimating the format of the note. It seems that mere experience without reflection does not lead to insight.

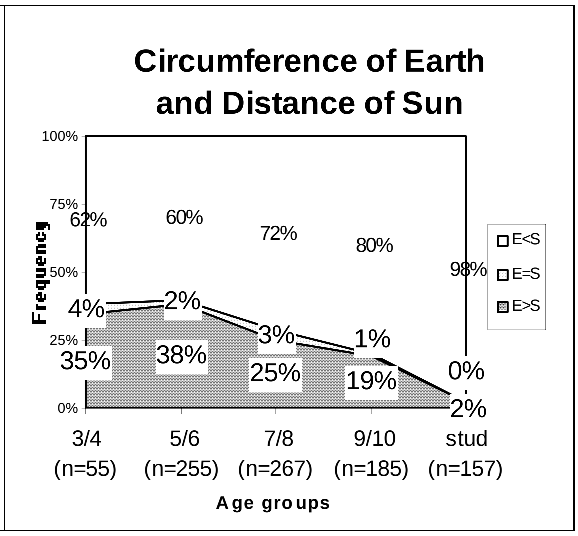
What is the effect if two systems of reference contradict each other? This is the case with coins, where sometimes size (diameter and weight) contradict value; the 10 Pf coin for example is 7.5% bigger in diameter (15.6% in surface) and 14.3% heavier than the more valuable 50 Pf coin. The diagrams show that for young children the value is overwhelming, but that with increasing age the sensual impression becomes more important than the monetary value. It is remarkable that about half of the mathematics students did not realize the greater weight of a 10 Pf coin.





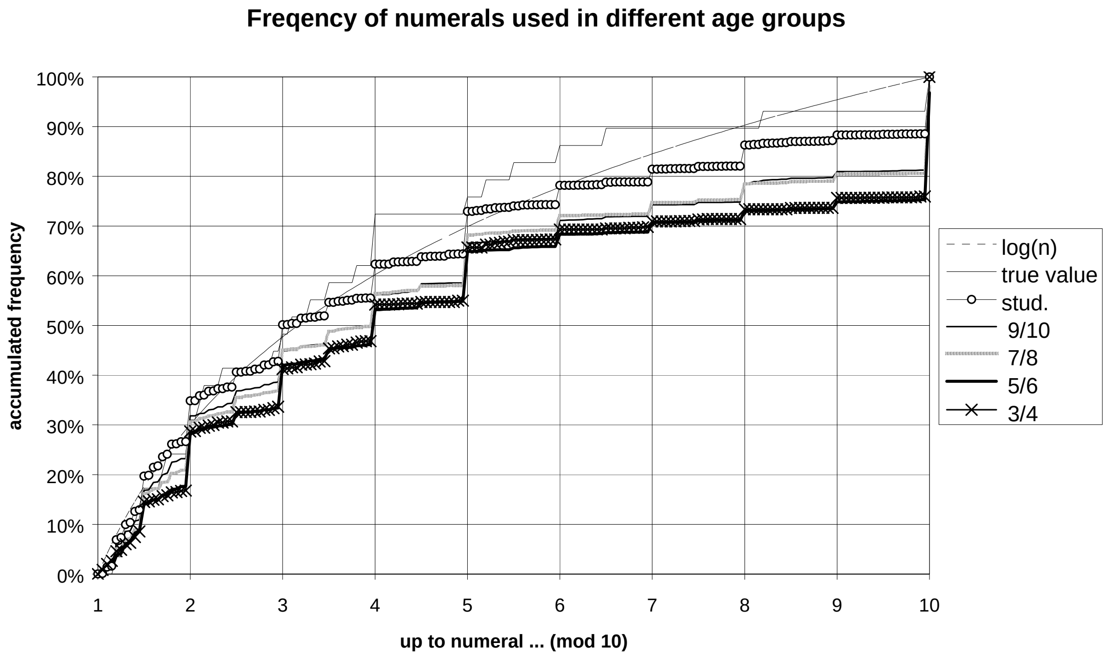
What can be found about the astronomical reference system of earth, moon and sun? Though the actual magnitudes are almost unknown, it could be asked if children are aware that the sun is more distant than the moon (which was discussed in public at least in connection with the eclipse of 11th August 1999)? Or, in the era of space travel, are children aware that the distance to the sun is bigger than the biggest distance on the earth (40 000 km)? The diagrams show that insight grows with increasing age, but that about $\frac{1}{4}$ of primary pupils think that the sun is not more distant than the moon as do 4% of the mathematics students. Also, the size of the earth is broadly overestimated in comparison to the distance of the sun (and about equally to the distance of the moon). Up to grade 6 about 40% of the pupils believe the sun to be closer to the earth than the length of the equator and at the end of secondary school 20% still believe this.





Which numbers are actively used in estimating?

A closer look at the numbers used in the estimation tests gives some unexpected insights into the numerical vocabulary which is available to persons in different age groups. When people estimate, they choose their numbers not only according to the supposed size of the object, but also according to the supply of numbers which are available in their brain. The numbers consist of two parts: their order of magnitude, which is represented by the number of places in the decimal system, and the numeral (i.e. the sequence of digits from left to right). If we neglect the order of magnitude and only look to the numerals used, we see that the use of numbers is according to Benford's law for the distribution of the first digit: the probability that a random decimal begins with the digit n is $\log_{10}(n+1) - \log_{10}(n)$ (Raimi). As the diagram shows, not only the true values but also the estimated values in all age groups approximately follow this law. A closer analysis shows that this pattern repeats inside every order of magnitude. That means if a person chooses a number between 1 and 10 from their numerical vocabulary, the frequency of the first digit of the number is not the same for all digits from 1 to 9, as could be expected if the underlying number concept was additive, but decreases from 1 to 9, which is more in accordance with a multiplicative number concept.



In the diagram another interesting pattern becomes visible. There are big thresholds in the course of the curves. That means, on the personal number lines, there are important threshold values which are chosen much more often than a value in the close neighbourhood. The most important threshold value for all age groups is 1 (or 10 etc., which in the diagram is seen at the end of the scale, as the height of the last step leading to 10), the next important ones are 2 and 5, then 1.5, 3 and 4, and then to a smaller extent 1.2, 2.5, 3.5, 6 and 8. The following table shows the relative frequency with which these numerals are used in the different age groups.

What are the most frequent numerals?

numerals used	**1**	1.2	1.5	**2**	2.5	3	3.5	4	**5**	6	8	*alto-gether*	**1; 2; 5**	1.5; 3; 4	1.2; 2.5; 3.5; 6; 8	*remaining numbers*
gr. 3/4	**24%**	2%	5%	**12%**	2%	7%	2%	7%	**11%**	2%	2%	*76%*	**46%**	20%	10%	*24%*
gr. 5/6	**21%**	2%	6%	**12%**	2%	9%	2%	7%	**10%**	3%	2%	*77%*	**44%**	22%	12%	*23%*
gr. 7/8	**17%**	3%	8%	**11%**	3%	8%	3%	7%	**11%**	3%	3%	*76%*	**39%**	23%	15%	*24%*
gr. 9/10	**14%**	4%	7%	**9%**	3%	7%	3%	6%	**9%**	4%	4%	*71%*	**33%**	20%	18%	*29%*
stud.	**11%**	5%	7%	**8%**	3%	7%	3%	7%	**9%**	4%	4%	*68%*	**28%**	21%	19%	*32%*

This shows something of the numeric alphabet which people use. The most important part consists of 11 one- and two-digit numbers. Young people use these 11 most frequent numerals for answering about 75% of the estimation questions; students a little bit less than 70%. The sequence of these numerals is not arithmetic with constant differences, but approximately geometric with constant ratios (1, $10^{1/11}$, $10^{2/11}$, $10^{3/11}$, $10^{4/11}$, $10^{5/11}$, $10^{6/11}$, $10^{7/11}$, $10^{8/11}$, $10^{9/11}$, $10^{10/11}$), which supports the idea of an underlying multiplicative number concept.

A closer look reveals different steps. *The first step* is the numerals **1, 2 and 5**, the benchmarks for doubling and halving (corresponding approximately to the sequence 1, $10^{1/3}$, $10^{2/3}$), which are used by primary children in almost every second answer, by secondary pupils in about 40% and by adult students in about 30% of the answers. In *the second step*, the gap between 1 and 5 is filled by the numerals 1.5 (or 15 etc.), 3 and 4, which are used in about 20% of the answers. In *the third step,* the numerals 1.2, 2.5, 3.5, 6 and 8 emerge, used in primary school in

10% of the answers, in secondary school in 15%, and by students in 20%. The use of the *remaining numerals* increases from $\frac{1}{4}$ of the estimation answers during school age to nearly $\frac{1}{3}$ of the answers after school.

The first step shows how important doubling and halving is for the development of number sense – operations which are mostly neglected after the first two school years in mathematics education. In the third step and in the use of the remaining numerals we notice a gradual refinement of the numerals actively used, probably due to the introduction of decimals in grade 6.

In all, we see how important round numbers are for the development of number sense. Additionally the diagram above gives an idea of what it means for consumers, especially young people, when companies fix prices a little bit below a round number. If you see the accumulated frequency curves as approximations to the subjective number value and compare them with the logarithmic curve, which would represent an accurate multiplicative number concept, you get a feeling of how a price like 9.80 DM matters for different age groups. For children in grade 3 to 6 the height of their frequency curve is about the same as the logarithmic value of 5.5, for young people in grade 7 to 10 it is about 6.5, and for adults it is a little bit less than 8.

Conclusions

Including numerical estimation and the estimation of measurements into daily practice in school maths could be an important way to improve number sense.

Estimation is important **for the pupils** because they can get immediate information about their success which is highly motivating, and because reflection about the actual measurements and about their personal reference systems can lead to improvement. Dealing with big numbers in estimation questions and discussing the underlying situations could be promising, especially for girls, in order to give them more experience in these areas and hence a better understanding of large numbers. It is also important because all children need these abilities quickly in real life. A better number sense could be developed by using doubling and halving as fundamental operations rather than as peripheral activities.

Including more estimation situations could be productive **for mathematics teachers**, as they can learn new things about their pupils and because they would be encouraged to find new approaches to counter the predominant innumeracy.

It is important **for researchers** to try to find out which situations lead to an additive (cardinal) reaction and in which situations a multiplicative (logarithmic) number concept is predominant. We also need to discover what cognitive coexistence of these two concepts implies. Where do contradictions arise? What are the consequences for the mathematics curriculum?

References

BENTON, S.E. 1986. A Summary of Research on Teaching and Learning Estimation. In: H.L.Schoen, M.J.Zweng (Eds.) Estimation and Mental Computation. 1986 Yearbook. NCTM. Reston, pp. 239-248.

BUTTERWORTH, B. 2000. The Mathematical Brain. Papermac London.

DEHAENE, S. 1998. The Number sense. How the Mind Creates Mathematics. Penguin London.

FUSON, K.C. 1992. Research on Whole Number Addition and Subtraction. In: D.A.Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning*. NCTM. New York, pp. 243-275.

GERSTER, H.D., SCHULTZ, R. 1998. Schwierigkeiten beim Erwerb mathematischer Konzepte im Anfangsunterricht. PH Freiburg.

GLENN, J.A. (Ed.) 1978. The Third R. Towards a Numerate Society. London.

HOFSTADTER, D.R. 1982. Metamagikum. In: Spektrum der Wissenschaft, Juli 1982, pp.8-13

MEIßNER, H. 1985. Vergleichen, Schätzen, Überschlagen. In: Beiträge zum Mathematikunterricht 1985. Franzbecker Bad Salzdetfurth, pp. 214-217.

PAULOS, J.A. 1996. A Mathematician Reads the Newspaper. Penguin London.

RAIMI, R.A. 1969. The Peculiar Distribution of First Digits. In: Scientific American (221) Dec. 1969, pp.109-120.

SOWDER, J. 1992. Estimation and Number Sense. In: D.A.Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning*. NCTM. New York, pp. 371-389.

WINTER, H. 1990. Aufklären durch Abschätzen. In: mathematik lehren Heft 39, pp.20-25

ZACHARIAS, J.R. 1974. The Importance of Quantitative Thinking. In: The National Elementary Principal (53) 2, pp.8-13.

Prof. Dr. Gustav Adolf Lörcher
Pädagogische Hochschule Freiburg
Kunzenweg 21
79117 Freiburg
E-mail: GALoercher@web.de