Klaus Hasemann, Hannover

EARLY NUMERACY - RESULTS OF AN EMPIRICAL STUDY WITH 5 TO 7 YEAR-OLD CHILDREN

Abstract.

The acquisition of mathematical skills can be viewed as a developmental process starting long before formal mathematics education begins. Based on the literature, Van de Rijt, Van Luit and Pennings distinguished eight components of early numeracy which were operationalized in a test for children aged 4 to 7 years. This test, the “Utrechtse Getalbegrip Toets (UGT)”, consists of 40 items, 5 for each component. In this paper, the results of three test runs in Germany with about 300 children each are presented and it is shown that, on average, children’s mathematical abilities at the beginning of school are rather high, but that there is also a wide range in the children’s performances.
In an additional project, 80 children were videotaped when solving some of the problems presented in the UGT; these problem-solving processes were analysed with regard to the strategies used. It is shown that the differences between the children’s results do not only derive from differences in their cognitive development but also from differences in their kinds of dealing with the mathematical objects.

The development of early numerical knowledge of children in kindergarten or at the beginning of school has been evaluated on many occasions by researchers all over the world (c.f., e.g., Gelman and Gallistel, 1978; Briars and Siegler, 1984; Fuson, 1988; Stern, 1999). Empirical studies of this kind are useful not only in order to achieve a correct picture of children’s abilities but also to recognize changes in this development; the relevance of these studies for mathematics teaching in primary schools is obvious.

From German studies, which have been carried out since the beginning of the eighties, it can be concluded that on the one hand the children’s abilities in the field of early numeracy are very often underestimated even by experts, but that on the other hand there is a wide range in the individual performances of school beginners (see, e.g., Schmidt and Weiser, 1982; Zur Oeveste, 1987; Selter, 1995; Grassmann *et al*., 1995). Most tests of this kind, however, focused on special aspects, for instance counting, or they included just a small number of test items. In 1994 Van de Rijt, Van Luit and Pennings published a new test, the “Utrechtse Getalbegrip Toets (UGT)” which was designed to measure *all aspects of early numeracy*. The test was developed and used in the Netherlands with more than 800 children. The main objective of test construction was to develop an unidimensional early numeracy scale on which both test item and subject performance levels could be represented in order to assess the inter- and intra-individual differences in the development of early numerical knowledge. As research has shown that arithmetic difficulties at school can partly be explained by an insufficient competence in early numeracy (Van de Rijt and Van Luit, 1998), in the Netherlands the test is also used as an early numerical competence test, especially for children with delayed development in this field.

The basis for the construction of the UGT was a review of studies on early numeracy development in young children which had been carried out in cognitive psychology as well as in mathematics education. This review provided a list of eight aspects of non-numerical and numerical quantity knowledge which contribute to early numeracy (see Van de Rijt, Van Luit, and Hasemann, 2000). In the test, the eight aspects are operationalized in eight components with five items each, hence the test consists of 40 items. It is available in two forms (test forms A and B). These eight components are:

1. Concepts of comparison i.e. the use of concepts in making a comparison between two non-equivalent cardinal, ordinal and measure situations. The subject has to demonstrate his/her knowledge of concepts in drawings of order relations. An example is: “Here you see Indians. Point to the Indian who has *less* feathers than the one you see here”. Gelman and Baillargeon (1983) show that 4 year-olds are able to compare non-equivalent situations using concepts such as low, lower, lowest, more and less, etc;

2. Classification (cf. Piaget, 1965). This component refers to the grouping of objects in a class on the basis of one or more features, for example “Look at these drawings. Point to the drawing without triangles”;

3. One-to-one correspondence (cf. Piaget, 1965). This component refers to what subjects understand about the one-to-one relationship of simultaneously presented objects. An example of this is: [The child has 15 blocks.] The experimenter shows a drawing representing two dice with the 5- and 6-pattern and asks: “Can you put as many blocks on the table as the numbers shown on the dice?” Overt and covert indicating acts (e.g., moving blocks, drawing lines, pointing) are necessary in responding to such one-to-one correspondence items;

4. Seriation (cf. Piaget, 1965). This component refers to dealing with discrete and ordered entities, for example: “Here you see drawings with apples. Show me the drawing where the apples are arranged from large to small”;

5. Using number words. This component refers to the use of number words that are in the number-word sequence up to twenty. Number words must be produced forward and backward. An example is: “Count further from 9 to 15: 6, 7, 8....”;

   Fuson (1988) reports that most children below age 3 ½ work on learning the sequence to ten while most children between 3 ½ and 4 ½ work on learning the sequence of number words between ten and twenty. However, many children between 4 ½ and 6 are still imperfect between fourteen and twenty in the sequence.

6. Structured counting. This component refers to the counting of objects in organized and disorganized arrangements with pointing acts. An example is: [The experimenter puts 20 blocks on the table in a disorganized arrangement.] The child is asked to count the blocks. [The child is allowed to point to the blocks with their finger or to move the blocks.] Fuson (1988) demonstrated that most of the children aged 5 ½ to 6 count correctly when pointing or moving is allowed;

7. Resultative counting. This component refers to accurate counting and last-word responding in which pointing acts are not permitted, for example: The experimenter puts 15 blocks down on the table in three rows of five with some space between them and asks: “How many blocks are there?”;

8. Applying general knowledge of numbers. This component refers to the application of number knowledge in real life situations which are represented in drawings. An example is: “You have nine marbles. You lose three of them. How many are left? Point to the picture with the correct number of marbles”.

(It should be remarked that although three components are derived from Piaget's work, their content is not fully comparable with the original Piagetian content. In this test the content is as much as possible based on numerical abilities, and counting can be used to solve items in, for example, a seriation or correspondence context.)

In our investigation we have used the UGT three times; the first time was when children were in the second half of their last year in kindergarten, the second time at the end of kindergarten and the third time in the middle of their first year at school. At the same time as our investigation the UGT was used by researchers in Finland, Greece and the U.K. The first test run involved 330 children (168 boys and 162 girls) within the Osnabrück area, the second run was taken by 306 of these 330 children (160 + 146) and the third by 292 children (148 + 144). Each test took 25 to 30 minutes. In Table 1 the results of the three test runs are given:

test run	means			standard deviations			range		age (means)	
	boys	girls	com-bined	boys	girls	com-bined	boys	girls	boys	girls
T1	23.4	24.0	23.7	6.8	7.3	7.1	7 - 38	5 - 39	6.2	6.2
T2	26.4	26.2	26.3	7.0	7.4	7.2	10 - 40	5 - 40	6.5	6.5
T3	32.9	32.7	32.8	4.7	5.2	5.0	12 - 40	14 - 40	7.1	7.1

Table 1: Test results of about 300 children in kindergarten and in their first year at school: Means, standard deviations, ranges and ages. The UGT includes 40 items.

There is a remarkable increase in correct answers from one test run to the next i.e. from 23.7 correct answers in the first test run T1 to 26.3 correct answers in the second (which took place at the end of kindergarten) and then to 32.8 correct answers in the middle of the first year at school. Thus in the last 3 ½ months of kindergarten there was an average increase of 2.6 correct answers. The third test run T3 was too easy for most of the children and this fact becomes very clear from the histograms in Figures 1, 2 and 3:

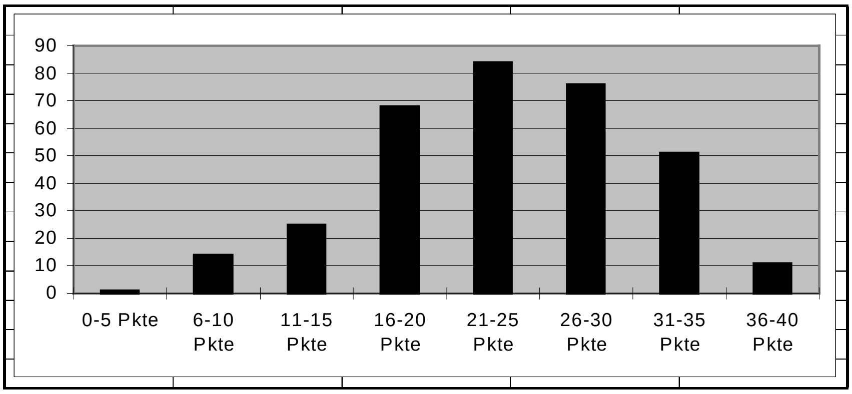

Figure 1: Number of correct answers in the first test run (N = 330)

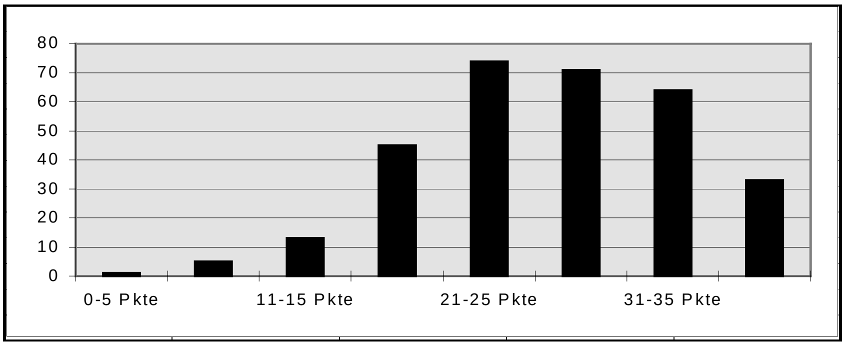

Figure 2: Number of correct answers in the second test run (N = 306)

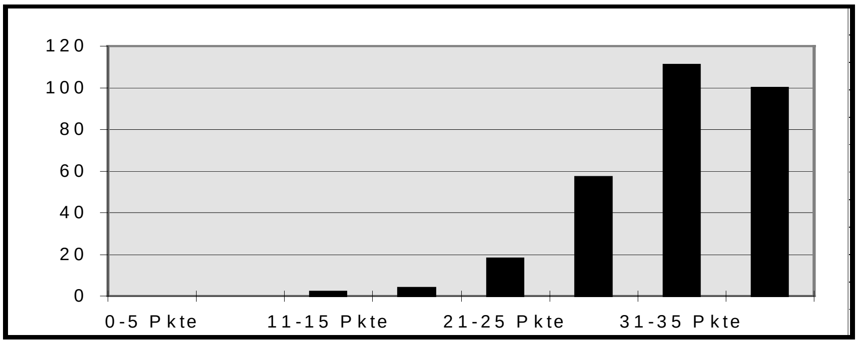

Figure 3: Number of correct answers in the third test run (N = 292)

Next the results for the eight components of the test are compared. In Table 2, however, not only are the mean values of correct answers of the whole samples at the three times of measurement (tests T1, T2 and T3) given but also, for the second test run (test T2), the mean values of three performance groups. Here “Q 1” indicates the best quarter of the whole sample, “Q 4” the weakest quarter and “Q 1-3” the whole sample minus the weakest quarter (each component consists of 5 items so that the maximum for each component and each test is 5).

component	number words	structured counting	resultative counting	knowledge of numbers
T1 (means)	2.6	2.9	2.2	2.8
T2 (means)	2.9	3.3.	2.5	3.1
T3 (means)	4.0	4.1	3.8	4.0
T2: Q 1	4.4	4.4	3.7	4.5
T2: Q 1-3	3.4	3.7	3.0	3.5
T2: Q 4	1.2	1.9	1.2	1.8

component	comparison	classification	correspondence	seriation
T1 (means)	4.5	3.5	3.3	2.1
T2 (means)	4.6	3.7	3.6	2.7
T3 (means)	4.8	4.1	4.4	3.4
T2: Q 1	4.9	4.3	4.6	4.4
T2: Q 1-3	4.8	3.9	3.9	3.2
T2: Q 4	4.0	3.0	2.7	1.1

Table 2: Mean values of correct answers for each component of the UGT for the whole group in the three test runs (tests T1, T2 and T3), and mean values of three performance groups (Q 1, Q 1-3, Q 4) for test T2.

Although the figures in Table 2 should not be overrated some quite interesting conclusions may be drawn namely:
1. the increase in the average number of correct answers from one test run to the next can be observed in each component in (more or less) the same way. The increase from T1 to T2, however, is above average in the component “seriation” which was the hardest part of the test;
2. regarding the results of the best quarter in the sample (Q 1) it seems that for these children the test is already too easy at the end of kindergarten (it should be remembered that in the UGT there are a lot of rather hard items on all aspects of numeracy including simple word problems or situations which ask for additions or subtractions, and that the test time was less than 30 minutes for the 40 items);
3. highly relevant for the beginning of mathematics teaching in school is the enormous range in the children’s performances in T2. In the five hardest components (seriation, number words, structured and resultative counting and knowlege of numbers) the difference between the mean values of the weakest group (Q 4) and the rest of the sample (Q 1-3) is about 2 items. This means that in each class of beginners there is a considerable group of children whose numerical competence is clearly weaker than that of the rest of the class (not to mention the differences between Q 4 and Q 1). This fact will unavoidably cause a lot of problems for the teacher. (It should be noted that in test T3, which took place in the middle of the first year at school, the differences between the groups Q 4 and Q 1-3 had diminished, but when interpreting this result it has to be taken into account that at that time the test was too easy for most of the children).

Regarding the wide range in the children’s performances with the UGT it might be asked where these differences come from; are they caused by differences in the individual’s cognitive development or are there also differences in the kind of thinking between the brighter and the weaker children? To answer this question we explored the strategies these children use to solve problems like those presented in the UGT. As it was not possible to trace the special kind of problem-solving strategies from the solutions of the 330 children mentioned above, we chose two more samples. These children worked out only some items from the UGT and they were given as much time as they wanted to solve the problems. The problem-solving processes were videotaped. In the first sample there were 42 first-graders at the very beginning of their school life and in the second sample 38 kindergarteners in the middle of their last year in kindergarten.

As a first step each sample worked out Raven’s test “Coloured Progressive Matrices” (CPM; cf. Becker, P.; Schaller, S. & Schmidtke, A., 1994) to give us additional data on the children’s cognitive development. From these data, together with their results with the items taken from the UGT for each sample, two sub-groups were formed. The first sub-group consisted of those 10 or 11 children who had been the brightest ones in both tests (in the CPM *and* in the items from the UGT), while the second sub-group comprised those 10 or 11 who had been the weakest ones in both tests. For these two sub-groups the (videotaped) problem-solving strategies with the items from the UGT were analysed and compared. In the following paragraphs we focus on the results of the first-graders with the results of the kindergarteners being quite similiar.

When selecting the items from the UGT for this special investigation, we chose, on the one hand, items from the “Piagetian” part of the UGT (i.e. from the components “classification”, “correspondence” and “seriation”) as it was to be expected that there would be differences in the children’s individual development. However as, for instance, Weinert and Helmke (1994, pp. 13-15) pointed out, the cognitive development cannot be described just by “universal” and “structural” aspects (as, for example, just by the use of the Piaget stages). Individual differences in the children’s kinds of thinking have also to be taken into account (see also Stern, 1999). Gray, Pitta, and Tall (1997, p. 117) went even further when they wrote: “It is our contention that different perceptions of (the) objects, whether mental or physical, are the heart of different cognitive styles that lead to success and failure in elementary arithmetic”. Following their ideas, we chose some items from the second part of the UGT which seemed promising in helping to elucidate children’s different kinds of thinking.

Initially we will discuss the children’s behaviour with some items of Piagetian type. There was, for instance, a classification item in which the children were asked to compare 12 drawings of apples with a model. In the drawings there were two distinguishing marks - leaves and a little worm. All 11 children of the brighter sub-group solved this problem correctly, but only 2 (out of 10) of the weaker sub-group did so. The mistakes in this group were all caused by the fact that these children just focused on *one mark* (either leaves or the worm), i.e. in Piaget’s terms they did a simple but not a multiple classification.

This is also true for seriation and correspondence. Here 9 children of the brighter, but only 2 of the weaker sub-group were able to do the seriation item in Figure 4, and a similiar result holds for the hardest correspondence item (see Figure 5).

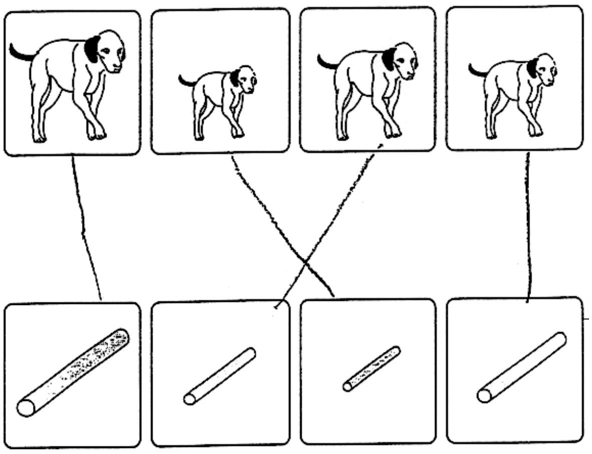

Figure 4: [Examiner gives the child a work paper and a pencil.] Here you see dogs. Each dog is going to get a stick. A big dog is going to get a big stick and a small dog is going to get a small stick. Can you draw lines from the dogs to the sticks they are going to get?

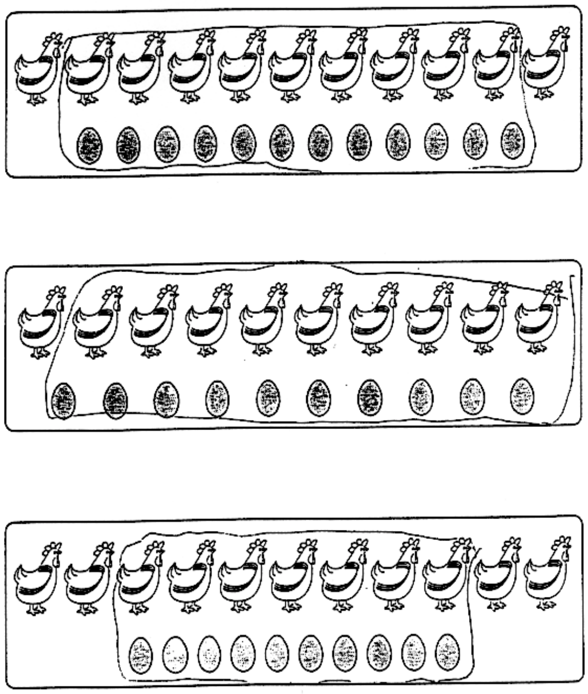

Figure 5: Here you see pictures of chickens and eggs. Can you find the picture where each chicken has laid one egg? You may draw lines.

On the other hand, if the children can establish the equivalence of two sets by counting, the task is rather easy even for the weaker ones.

It was to be expected that with items of the Piagetian type “weaker” and “brighter” children would behave differently. In our sample, however, there were also children from the weaker sub-group who solved the hardest classification or correspondence items correctly whereas some children from the brighter group failed with these items.

In addition to differences in the cognitive development which may be described in terms of Piaget’s theory, children obviously differ in their abilities to solve problems related to numbers and quantities (see Gray, Pitta and Tell’s statement cited above). With special regard to preferences, we tried to describe these differences by analyzing the problem-solving strategies which were used by the children of the two sub-groups. The item in Figure 6, for instance, was solved correctly by all 11 children of the brighter sub-group and also by 8 of the 10 children of the weaker sub-group. There were, however, very big differences in their strategies. The dominant strategy of the “weaker” children was counting whereas 10 of the 11 “brighter” children used so-called “heuristic strategies” i.e. they either recognized in the drawings the dice-like pattern of six points and as a result of this they found the drawing with seven points, or they pointed to this drawing and made sure that they had seen 3 + 1 + 3 points.

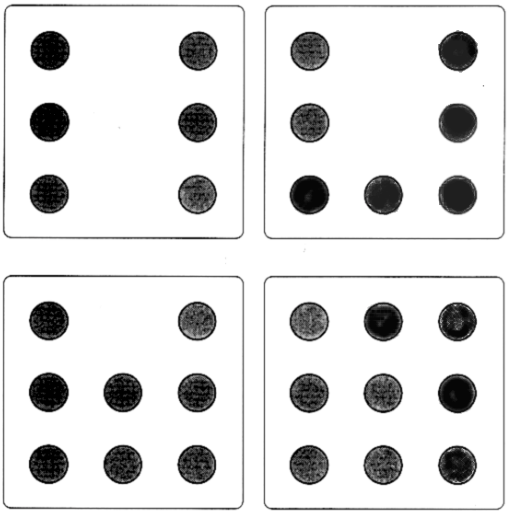

Figure 6: [Examiner shows the picture to the child.] Point to the picture with seven dots.

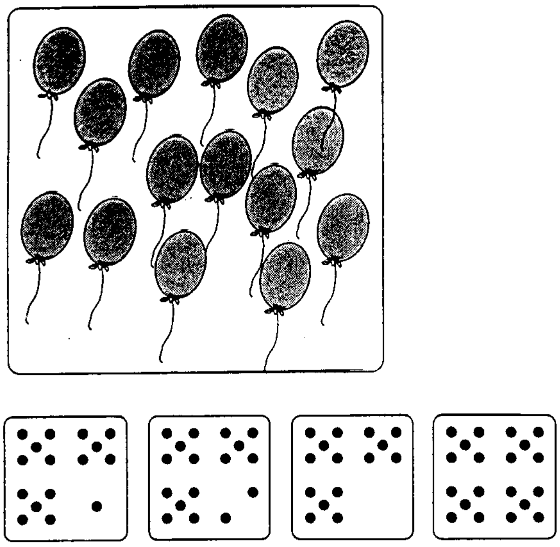

Figure 7: Here you see fifteen balloons. [Examiner points to the balloons in the picture on the top of the page.] Point to the square where there are as many dots as there are balloons.

Another example of this kind is presented in Figure 7. This item was much harder than the previous one (seven correct answers in the brighter sup-group and none in the weaker). Once again we find three (brighter) children who used calculations (“5 + 5 + 5” or “3 times 5”) whereas *all* children in the weaker group pointed to the drawing with 20 points “because this has more (or a lot of) points”.

The hardest item in the whole test was:
[Examiner lays down 5 blocks on the table.] Here are five blocks. I am going to push them under my hand. [Examiner pushes the blocks under a hand. Then examiner pushes another 7 blocks, which are shown to the child, also under the hand.] I add seven blocks. How many blocks are there under my hand?
Three children of the brighter sub-group used calculations: “7 + 3 + 2” (twice) or even “5 + 7 = 6 + 6 = 12” whereas the others counted up from 5. Eight of the children in the weaker group, however, tried to make a guess, e.g. 5, 17, 18, or 80.

In the counting items we also found remarkable differences between the children. For example, in the item: [Examiner lays down 20 blocks on the table in a heap with small distances between the blocks.] Count these blocks. [The child is allowed to point to the blocks or to push them aside while counting], there were nine correct solutions in the brighter group, but only four in the weaker. It should also be noted that in the weaker group the mistakes mainly came from the fact that while counting most of these children did *not* point to the blocks or push them aside although they were allowed to do so.

From these differences in the children’s problem-solving strategies we can conclude that there are, in fact, differences in their kinds of thinking. Most of the children in the brighter sub-group were flexible and confident in using the sequence of number words, they were able to count up and down, they could do simple calculations and used, if possible, patterns. They could also select effective and reliable procedures to solve the problems. On the other hand, most of the weaker group could add only by counting up, had no techniques to prove results and very often trusted in their visual impression.

Some conclusions for the teaching at the very beginning of school can be drawn. Teachers as well as curriculum designers have to take into account that the early numerical abilities of most of the children are much higher then adults (and even experts) believe. However there is also a wide range in these abilities and hence teachers at the beginning of school run the risk of demanding too little from most of the children but, at the same time, demanding too much from the weakest. The differences between the children do not only come from differences in their cognitive development but also from differences in their kind of dealing with the mathematical objects. This analysis leads us to recommend a differentiated and - as far as possible and socially acceptable - individualzed instruction style in school which supports the intelligent use of heuristic strategies for the brighter children while at the same time helping the weaker children to adopt these strategies.

Regarding this last point, we agree with Wittmann and Müller (1995) who came to the conclusion that this does *not* mean that mathematical instruction in school must be based on problem solving or depends on the brighter pupils' creative ideas: "This is a big misunderstanding ... ; however, even the weaker pupils cannot learn effectively if they are not active for themselves and do not contribute to classroom activities. Especially for these pupils it is extremely important to be shown the relationships between mathematical topics as they, different to the brighter children, cannot establish the basic orientations by themselves which are necessary for learning". (1995, p. 165; translation K.H.)

In a new project we are working with younger children (aged 4 to 6) who are asked to solve problems similiar to those from the UGT. The target is to find out in which situations, realistic or mathematical ones, these children are likely to find, by themselves, the relations between the characteristics of the situations and the mathematical objects, activities or stuctures which are included in these situations. This experiment might help us to connect our findings about children's differences in dealing with mathematical objects with Gray, Pitta and Tall's ideas on the different mental representations of the mathematical objects used by children.

References

BECKER, P.; SCHALLER, S., AND SCHMIDTKE, A: 1994, *Coloured Progressive Matrices, Testmappe*, Beltz, Weinheim.

BRIARS, D.J. AND SIEGLER, R.S: 1984, 'A featural analysis of preschoolers' counting knowledge', *Developmental Psychology,* 20, 607-618.

FUSON, K.C.: 1988, *Children's counting and concepts of numbers*, Springer, New York.

GELMAN, R. AND BAILLARGEON, R.: 1983, 'A review of some Piagetian concepts', in P.H. Mussen (ed.), *Handbook of child psychology, Vol. III*, Wiley, New York, pp. 167-230.

GELMAN, R. AND GALLISTEL, C.R.: 1978, *The child's understanding of number*, Harvard University Press, Cambridge.

GRAY, E., PITTA, D, and TALL, D: 1997, 'The nature of the object as an integral component of numerical processes', in *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education*, 14 - 19 July 1997, Lahti, Vol. 1, 115-130.

GRASSMANN, E. et al.: 1995, 'Arithmetische Kompetenz von Schulanfängern - Schlußfolgerungen für die Gestaltung des Anfangsunterrichts', *Sachunterricht und Mathematik in der Primarstufe,* 23, 314-321.

PIAGET, J.: 1965, *The child's conception of number*, Norton, New York.

SCHMIDT, S. and WEISER, W.: 1982, 'Zählen und Zahlverständnis von Schulanfängern: Zählen und der kardinale Aspekt', *Journal für Mathematik-Didaktik*, 3, 227-263.

SELTER, C.: 1995, 'Die Fiktivität der „Stunde Null" im arithmetischen Anfangsunterricht' *Mathematische Unterrichtspraxis, 16(2)*, 11-19.

STERN, E.: 1992, 'Lern- und Denkstrategien in Schule und Unterriccht', in H. Mandl and H.F. Friedrich (eds.), *Lern- und Denkstrategien*, Hogrefe, Göttingen, 99-123.

STERN, E.: 1999, 'Development of mathematical competencies', in F.E. Weinert and W. Schneider (eds.), *Individual development from 3 - 12: Findings from the Munich Longitudinal Study.* Cambridge University Press.

VAN DE RIJT, B.A.M. and VAN LUIT, J.E.H.: 1998, 'Effectiveness of the AEM program for teaching children early mathematics', *Instructional Science,* 26, 337-358.

Van Luit, J.E.H., Van de Rijt, B.A.M., and Pennings, A.H.: 1994, *Utrechtse Getalbegrip Toets* [Early Numeracy Test], Graviant, Doetinchem.

VAN LUIT, J.E.H., VAN DE RIJT, B.A.M. and HASEMANN. K.: 2000, 'Zur Messung der frühen Zahlbegriffsentwicklung'. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 32(1), 14-24.

WEINERT, F.E. and HELMKE, A.: 1994, 'Wie bereichsspezifisch verläuft die kognitive Entwicklung?', in *Lern- und Lehrforschung. LLF Berichte Nr. 9.* Interdisziplinäres Zentrum für Lern- und Lehrforschung, Universität Potsdam, 13-35.

WITTMANN, E.C. and MÜLLER, G.: 1995, *Handbuch produktiver Rechenübungen*, Vol. 1, Klett, Stuttgart.

ZUR OEVESTE, H.: 1987, *Kognitive Entwicklung im Vor- und Grundschulalter*, Verlag für Psychologie, Göttingen.

Klaus Hasemann
Universität Hannover
Institut für Didaktik der Mathematik
Bismarckstr. 2
D-30173 Hannover