Realistic Math Education

J. de Lange Jzn
OW & OC, R.U. Utrecht

Samenvatting

In deze bijdrage aan het OW & OC-symposium op 2 mei 1986 schetst Jan de Lange aspecten van het Hewet project als representant voor de zogenaamde realistische benadering van het wiskunde-onderwijs.
Via Wiskobas heeft deze benadering veel invloed verworven in het basis-onderwijs. Maar ook op een hoger niveau zijn deze principes gerealiseerd in het experimentele Hewet-materiaal. Gedeelten van dit artikel zijn reeds eerder in de Nieuwe Wiskrant gepubliceerd.

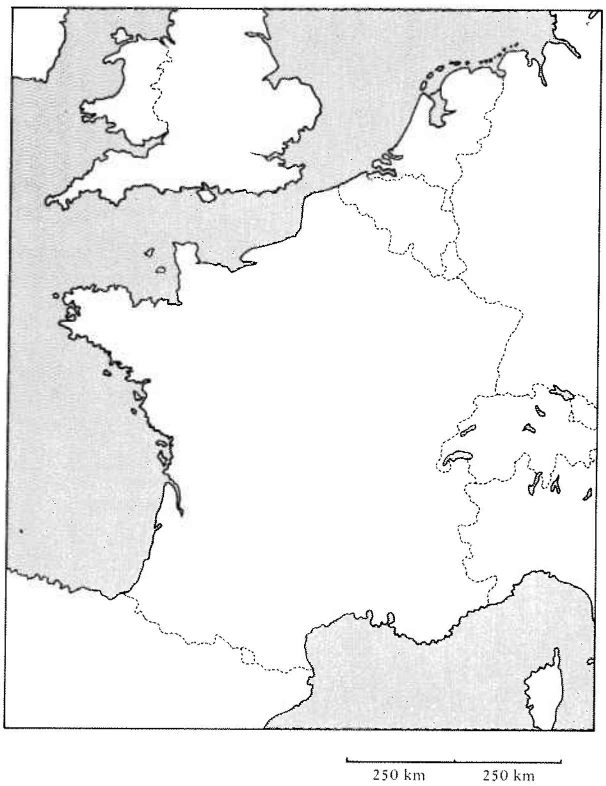
Introduction

The Olympics of 1984 gave rise to the following article in the Dutch newspaper *'De Volkskrant'* of 13.8.1984.
The problem discussed concerned the number of goldmedals related to the size of a country:

> *"It takes a lot of arithmetic, so let's restrict ourselves to the Netherlands. That country has 14 million inhabitants, and the U.S.A. more than 3 billion, two hundred times as much.*
> *The area of the Netherlands is, roughly speaking, 40.000 square meters, against 33.000 square kilometers for the U.S.A.; almost a thousand times as much.*
> *Considering all these facts leads to a population coefficient for the Netherlands that is one-fifth of that of the U.S.A."*

This clip was given to many teacher-students in the Netherlands, preparing for a career as elementary school-teacher.
The question was: "How many mistakes can you find in this clip?"
All these students were properly trained in the metric system. They know exactly that:
$1 \text{ km}^2 = 1.000.000 \text{ m}^2$.
They have made numerous exercises of this kind.
The question is, however, does all this has a meaning for the students?
Obviously, looking at the results, not at all. Less than 10% of the students notices only 1 mistake out of – at least – four mistakes. Nobody notices that 40.000 m² is even for a tiny country like the Netherlands too modest. The students know the metric system, but their knowledge is of no practical use because of the lack of real reference-points.

More than ten years ago *Wiskobas*-predecessor of OW & OC – developed materials like:



- Make an estimation how much larger France is compared to the Netherlands.
- The Netherlands has 14 million inhabitants.
  Make an estimation of the number of inhabitants of France.

Here one sees how one could start with real problems as an introduction to the metric system. Children first make estimations, get a feeling about comparing areas, before moving on to the more formal aspects of areas and the metric system.
The times are changing however. The influence of Wiskobas is clearly visible in the presentday books. Research carried out by Streefland [1] and Van den Brink [2] leads to new subcurricula with very promising results.
Treffers [3] is the stimulating factor behind a *National Plan* to reform arithmetic and mathematics-education. Reactions to this plan from more than 200 experts lead to the conclusion that the way is free towards more realistic mathematics.

Mathematizing; an example

Five years OW & OC means also five years HEWET. This project was carried out by OW & OC in order to develop a new curriculum for upper secondary education for non-mathematical-majors. [4].
The project started in 1981 with the development of experimental material, and try-outs in two schools. Later on 10 and 40 more schools joined the experiments.
That *using* your mathematics is something different from *knowing* your mathematics became – once more – clear at teacher training courses that were given to the teachers of these schools, in order to prepare them for the new course.
The example shows also a special and characteristic feature of mathematics A. It is the process (of mathematization) that we are interested in, not the product. In this case the answer, the product, is given beforehand.
We agree, with Lesh, Landau, Hamilton [5] that many of the problems should be designed so that the critical solution stages would be 'nonanswer giving' stages. For example, in many realistic problem situations, problem formulation or trial answer refinement are crucial in the solution process. Thus, in many cases the goal is nót to produce a numerical 'answer', instead it was to make nonmathematical decisions, comparisons or evaluations using mathematics as a tool.
Let us now turn to our problem. We start with a page from the book *'Rats'* by 't Hart [6]:

Growth of rat populations

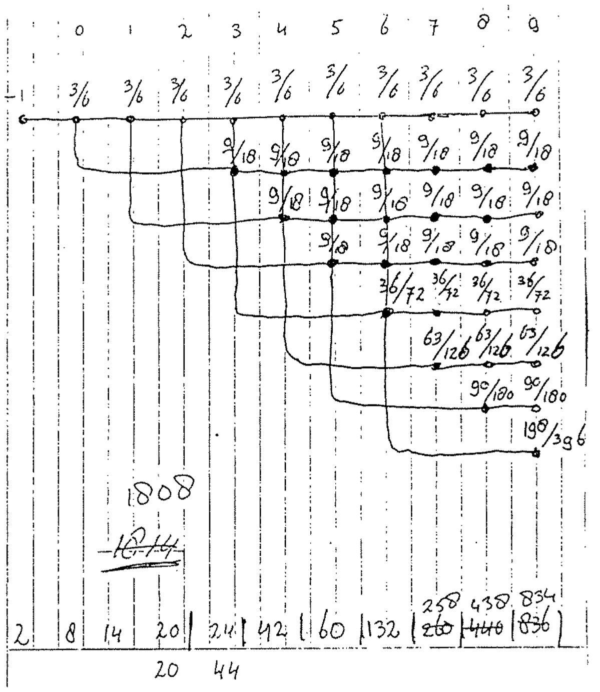
As regards the progeny of one pair of rats during one year the numbers given vary considerably. In the next chapter I shall discuss the scanty information supplied by research into the fertility of rats in nature, but at this point it might be interesting to estimate the number of offspring produced by one pair under ideal conditions. My estimate will be based on the following data. The average number of young produced at a birth is six; three out of those six are females. *The period of gestation is twenty one days; lactation also lasts twenty one days. However, a female may already conceive again during lactation, she may even conceive again at the very day she has dropped her young. To simplify matters, let* the number of days between one litter and the next be fourty. *If then a female drops six young on the first of january, she will be able to produce another six forty days later.* The female from the first litter *of six will be able to produce offspring themselves* after a hundred and twenty days. *Assuming there will always be three females in every litter of six, the total number of rats will be 1808 by the next first of january, the original pair included. This number is of course entire ficticious. There will be deaths; mothers may reject their young; sometimes females are not on heat for a long time. Nevertheless this number gives us some idea of the host of rats that may come into being in one single year.*

From: "Rats" by
Maarten 't Hart.

The question to teachers and students as well is:

"Is the conclusion that there will be 1808 rats at the end of the year, correct?"

The teachers concerned were teachers that took part in the teacher training courses that were given as part of the HEWET-project. These teachers had no previous experience in teaching mathematics A, other than from the course. The students, most of the time, encounter this problem in their last year.
The general trend among teachers was that only a very few were able to 'proof' that the answer 1808 is correct. By a very few we mean less than 20%, at best, in a time span of 20 to 30 minutes.
For the students it is impossible to give a general idea about their ability of solving the problem. But from information from some schools one is tempted to say that certain students did very well on this problem. Results depend of course also on the conditions: in the classroom, with a limited amount of time, students find it very difficult to solve, or even to schematize the problem. But given ample time, for instance by giving the problem as homework, some fine results came out.



The first impression, when observing teachers as well as students, is that they are overwhelmed by the amount of information, all relevant. So the main difficulty was to select and organize the information that is 'useful' in order to find a schema, as the next phase of mathematization.
Two students came with an original way of solving the problem. The first one is a schema that surprises by its simplicity. (See the preceding page).

The entries in the schema are the new borns only. 36/72 means: 72 newborns of which 36 are women. At the same time it is clear where those 72 young ones are coming from. Actually the schema is so simple that one wonders why this solution was not earlier found.
Another special solution – and one can see from the solution that it is not a teachers-solution – is from a girl that wanted to track down at any moment where the babies were coming from. This way of notation – her own invention – was perfectly clear to her, but she was unable to explain her way of writing down her solution in such a way that other students were convinced. Let us first present her solution:

1 worp : 6 jongen : 3 vrouw + 3 man
draagtijd 21 dagen + 21 dagen zogen
periode tussen 2 bevallingen = 40 dagen
jong vrouwtje na 120 dagen (3 x 40 dagen) jongen voortbrengen.

t = -1 man + vrouw (0) = 2

→ 3 meisjes x6 = 18 jongen

t = 0 2 + 6 jongen (0)(1) = 8 (0) = van 0 (1) = eigen nr.

40 t = 1 8 + 6 jongen (0)(2) = 14

80 t = 2 14 + 6 jongen (0)(3) = 20 → 18 → 9 meisjes x6 = 54 jongen

120 t = 3 20 + 6 jongen (0)(4) + (3x6 jongen (1)(5)) = 44

160 t = 4 44 + 6(0)(6) + 18(1)(7) + 18(2)(8) = 86

200 t = 5 86 + 6(0)(9) + 18(1)(10) + 18(2)(11) + 18(3)(12) = 146

240 t = 6 146 + $\underbrace{6(0)(13) + 18(1)(14) + 18(2)(15) + 18(3)(16) + 18(4)(17)}_{132}$

27 meisjes x6 = 162 jongen 54(5)(18) = 278

280 t = 7 278 + $\underbrace{132 + 18(6)(19) + 54(7)(20) + 54(8)(21)}_{258}$ = 536

320 t = 8 536 + $\underbrace{258 + 18(9)(22) + 54(10)(23) + 54(11)(24) + 54(12)(35)}_{438}$ = 974

360 t = 9 974 + 438 + 18(13)(26) + 54(14)(27) + 54(15)(28)

54(16)(29) + 54(17)(30) + 162(18)(31) = 1808

She explains her way of notation as follows:
in (0)(1)
(0) means: from 0
(1) is own number

What she actually has done is putting labels to any new birth. The first number of a couple indicates the parents(number) of the newborns, the second number indicates the number (ordinal) of the birth.
So at t = 4 we see:

44 + 6(0)(6) = 18(1)(7) + 18(2)(8) = 86

This means:

44 from t = 3

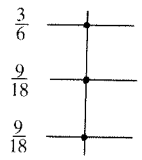
plus 6 newborns from the original parents which were labeled (0); this 6 come from the 6th birth this year;

plus 18 newborns from parents (1); being the 7th birth this year;

plus 18 newborns from parents (2); being the 8th birth.

Actual her mental construction and schematization is the same as the one from the previous solution:
Her
t = 4: (44+) 6(0)(6) + 18(1)(7) + 18(2)(8)
is similar to:

t = 4



Both solutions show a high ability to mathematize the problem, one in a more visual way.
Let us look now at one of the teachers-solutions showing already some shortcuts in the schemas that were presented as their solution. The first one:

t:	totaal aantal ratten
0	2 + 6 = 8
1	2 + 2·6 = 14
2	2 + 3·6 = 20
3	2 + 4·6 + 1·18 = 44
4	2 + 5·6 + (1+2)·18 = 86
5	2 + 6·6 + (1+2+3)·18 = 146
6	2 + 7·6 + (1+2+3+4)·18 + 54 = 278
7	2 + 8·6 + (1+ ... +5)·18 + (1+3)·54 = 536
8	2 + 9·6 + (1+ ... +6)·18 + (1+3+6)·54 = 974
9	2 + 10·6 + (1+ ... +7)·18 + (1+3+6+10)·54 + 162 = 1808

Progressive schematizing lead another teacher to the following solution:

t	-1	0	1	2	3	4	5	6	7	8	9
N	2	6	6	6	½·8·6= 24	½·14·6= 44	½·20·6= 60	½·44·6= 132	½·86·6= 258	½·146·6= 438	½·278·6= 834
T	2	8	14	20	44	86	146	278	536	974	1808

These two solutions are already so strict organized and schematized that the stage of presenting the relation in a formula is almost reached.
The problem, as it was stated, was solved. But especially teachers were not satisfied before having found a formule, or some kind of generalisation. And the booklet request children to do the same.
The next – teachers – solution leads almost directly to the formula (this teacher had already read an article [7] on the problem):

A(−1) = 2
A(0) = A(−1) + 6
A(1) = A(0) + 6
A(2) = A(1) + 6
A(3) = A(2) + 4.6
A(4) = A(3) + 7.6

The only missing link is now to replace
A(3) = A(2) + 4.6

by

$A(3) = A(2) + \frac{1}{2}.6.A(0)$

which gives:

$A_{n+3} = A_{n+2} + 3 \cdot a_n$

$A_{-1} = 2;\ A_0 = 8;\ A_1 = 14.$

We now have found a formula representing the growth.
Some teachers – although very few – had the feeling that another way of mathematizing and modelling was possible if not preferable.
This feeling was based on the fact that the period of 40 days as well as the one of 120 days plays such an important role, giving a suggestion that a partition of the population of rats into three age groups would be a 'natural' thing to do. Would it be possible to use population-projection-matrices (Leslie) in this case. Was this problem no isomorphic to other ones encountered with beetles and hooded seals?

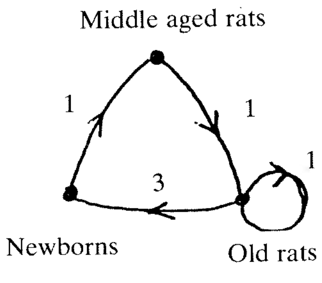
The start was a partition in newborns (0–40 days) and old ones (>80 days). Starting from the schema of the second teacher we get:

time	–1	0	1	2	3
newborns	0	6	6	6	24
middle	0	0	6	6	6
old	2	2	2	8	14

with the structure:

time	–1	0	1	2	3
newborns	0	6	6	6	24
middle	0 +3	0 +3	6	6 +3	6
old	2	2	2	8	14

which, in turn, can be represented in a graph:



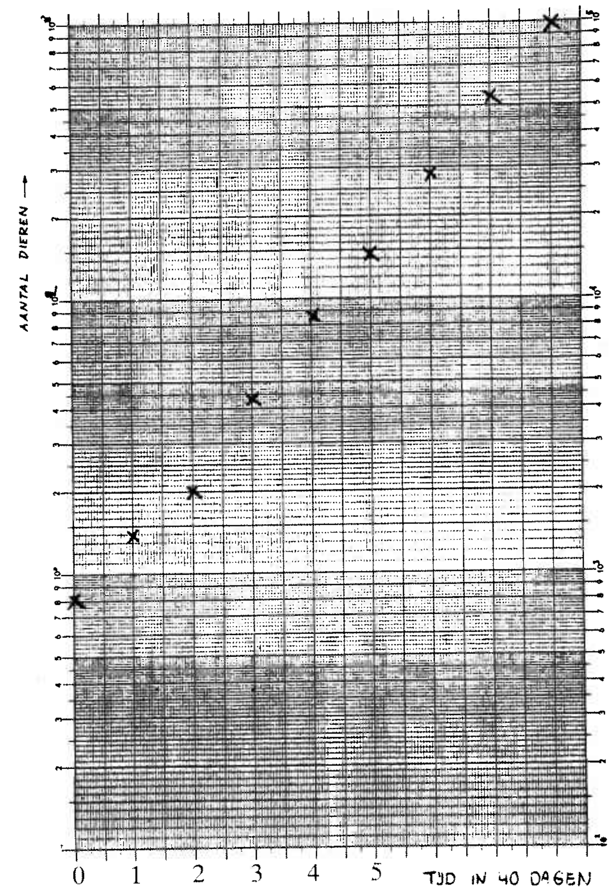
The meaning of the graph:
a. Newborns are certain to promote to the middle group (100% = 1).
b. The middle aged rats are certain to promote to the old group.
c. The old rats are certain not to die (!).
d. The old rats reproduce 3 females per period.

So the "1" 's represent surviving-probabilities, and the "3" the reproduction rate.

This visual representation can easily be translated to a matrix:

$$\begin{array}{cccc} & & \text{from} & \\ \text{to} & \begin{pmatrix} 0 & 0 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} & & \begin{array}{c} \text{n} \\ \text{m} \\ \text{old} \end{array} \\ & \text{n} \quad \text{m} \quad \text{old} & & \end{array}$$

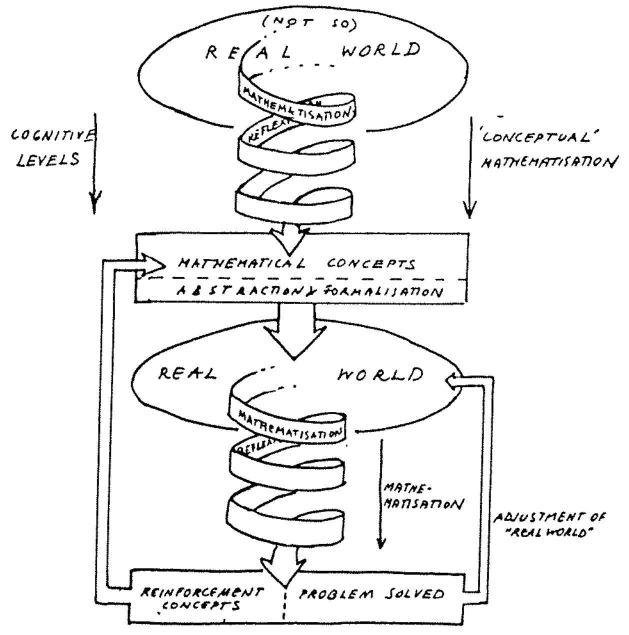
Our starting population of 2 will give us 1808 as the total population after 1 year. Of course the calculation can be done with aid of the computer. The advantage of this solution is that one has visible at all times the structure of the population. So one might ask for instance if the population-pyramid is stable.
Another question that might rise: "Is the growth of the rats an exponential growth?"
The simplest solution – certainly for the students – is to check whether the growthfactor $\frac{A_n}{A_{n-1}}$ is a constant.
Or, equally simple, one might draw the graph on logarithmic graph paper.
The graph for the first year on logarithmic paper leads to the conclusion that indeed we have at hand a case of exponential growth:



Further mathematization including Pascals triangle and more advanced mathematics is possible. But for the moment we leave our problem at this stage.

Mathematatization

Analysing the experimental material developed for mathematics-A is a very complex task. Not only does mathematization play a major role in solving real world problems, as indicated by our example, but also the start of any sub-curriculum takes place in some real world situation.
This real world situation or problem is explored intuitively in the first place, resulting in mathematizing this real situation. This means organizing and structuring the problem, to try to identificate the mathematical aspects of the problem, to discover regularities. This initial phenomenological exploration with a strong intuitive component will, in this stage, lead to the development, discovery or (re)invention of mathematical concepts.
Depending on factors as interaction between students, between student and teachers, the social environment of the student, the ability to formalise and abstract, the students will sooner or later extract the mathematical concepts from the real situation. Keywords of this phase of *conceptual* mathematization are intuitive exploration, reinvention, construction, strong interaction between the participants in the learningprocess.
At the same time reflexion on the process of mathematization is essential. The time at which reflexion is desired depends on the way a student proceeds from the first exploration to the formal/abstract phase.
The next phase recognisable in the material is the description of the desired and resulting mathematical concepts, followed by a more strict and formal definition.
Until now the context has served in the first place as a source for concept-building. In the next stage the mathematical concepts are being applied. The mathematization process now is comparable to the classical 'modelling and applications' description.
To put in another way: in our first process of mathematization we developed our tools; in the second stage we use our tools.



Of course this is a rather extreme view: by applying the concepts to new problems one of the main results is that the concepts are being reinforced. At the other hand the problem solved will have its effects of the students view at the real world: he adjusts his picture of the real world.
This very global sketch of the activities in the math-A curriculum are represented in the above schema.

Mathematization in Mathematics A

Mathematizing is an organizing and structuring activity according to which acquired knowledge and skills are used to discover unknown regularities, relations and structures. [8]. We may distinguish two components in mathematization, according to Treffers and Goffree [9]. First we can identify that part of mathematization that is aiming at transferring the problem to a mathematically stated problem. Via schematizing and visualizing we try to discover regularities and relations, for which it is necessary to identify the specific mathematics in a general context.

Activities with a strong horizontal component are:
- to identify the specific mathematics in a general context;
- schematizing;
- to formulate and visualize a problem in different ways;
- to discover relations;
- to discover regularities;
- to recognize isomorphic aspects in different problems;
- transfer a real world problem to a mathematical problem;
- transfer a real world problem to a known mathematical model.

As soon as the problem has been transformed to a more or less mathematical problem this problem can be attacked and treated with mathematical tools; the mathematical processing and reburfishing of the real world problem transformed into the mathematics.
Some articles that have a strong vertical component are:
- to represent a relation in a formula;
- to 'proof' a regularity;
- to refine and adjust models;
- to use different models;
- to combine and integrate models;
- to formulate a new mathematical concept;
- generalization.

Generalization may be seen as the top level of vertical mathematization. We mean, with Hilton [10], that when we are reasoning within the mathematical model we may feel compelled to construct a new mathematical model which embeds our original model in a more abstract conceptual way. Furthermore in learning the concrete real world problem we are improving the problem solving mentality of the students.

To divide the clusters of activities of mathematization in two distinct components is rather arbitrary, at least. As we have shown in our example, the two components are always intertwined. But the bipartition in a descriptive sense can be useful, not only to describe mathematization more clearly in concrete examples, but also to discriminate between different methodologies.

Nothing new

"To know mathematics means to be able to do mathematics: to use mathematical language with some fluency, to do problems, to criticize arguments, to find proofs and, what may be the most important activity, to recognize a mathematical concept in, or to extract it from, a given concrete situation. Therefore, to introduce new concepts without a sufficient background of concrete facts, to introduce unifying concepts where there is no experience to unify, or to harp on the introduced concepts without concrete applications which would challenge the students, is worth than useless: premature formalisation may lead to sterility; premature introduction of abstractions meets resistance especially from critical minds who, before accepting an abstraction, wish to know why it is relevant and how it could be used."

These words are taken from a memorandum that was published in 1962 in the Mathematics Teacher and the American Mathematical Monthly, and signed by distinguished mathematicians like Bers, Birkhoff, Courant, Coxeter, Kline, Morse, Pollak and Polya. [11]
This plea seems to support our process of conceptual mathematization, which is one of the characteristics for Realistic Mathematics Education.
Other characteristics are:
- the large-mental-contribution students make to the course;
- the interactive character of the learning process;
- the intertwining of learning strands.

Realistic mathematics

The horizontal and vertical components of mathematization allow us to discriminate between the four important ways of thought which can be internationally be distinguished in textbook series [8]:

- the empiristic view;
- the realistic view;
- the structuralistic view;
- the mechanistic view.

The empiristic approach is characterized by much attention for 'environmental' activities, more than for 'mental' operations. Horizontal mathematization is very strong, but less attention is paid to vertical mathematization.
In the realistic view – like in Math A – much attention is given to both horizontal and vertical mathematization: the process starts with an intuitive exploration resulting in the development of mathematical concepts.
In the structuralistic view much attention is given to insight in the structure of mathematics, mostly band on Set-theory as a direct result of the Royaumont-conference in 1959. [12]

The mechanistic view is directed at rules. The rules are given to the students, they verify and apply them to problems that are similar to examples given beforehand.
It may be useful to compare the four approaches in terms of horizontal and vertical mathematization.
The following scheme can be made:

	hor.math	vert.math.
empiristic	+	−
realistic	+	+
structuralistic	−	+
mechanistic	−	−

The + signs mean that much attention is given to either horizontal or vertical mathematization, and the − sign no or little mathematization.
As we described before, in the empiristic methods there is lot of mathematization but most of it is of a horizontal nature. In the realistic nature, considerable attention is paid to both horizontal and vertical mathematization. In the structuralistic approach almost all attention is described to vertical mathematization.
In the mechanistic methodes there is hardly any mathematization at all.

The experiences

The new curriculum was introduces after experiments at two schools (1981), ten schools (1983), forty schools (1984) and finally nationwide (1985) [13].

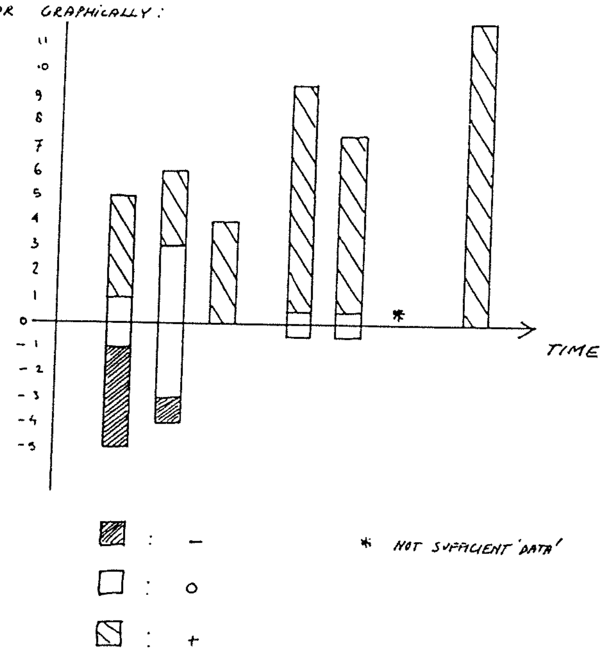
At the two schools the observations were at micro-level: almost all lessons in mathematics-A were observed by the members of the *Hewet*-team. The distance between the team and the actual school-practice was at a minimum. It was clear from the start that in the next stage this distance had to be much larger. The teachers of the ten schools followed a teacher training course given by the members of the *Hewet*-team. This was a facility not available for the teachers of the two schools. During this course soundings took place to find out how experiences were in the pre-experimental fourth class. This soundings took place in December 1982, March 1983 and May 1983. It was at this occasion that the discussion with the students of the two schools took place. During the next years the contact between the *Hewet*-team and ten schools consisted of a number of meetings to discuss the experiences. Of all this meetings internal written reports exist [14] from which we have extracted a general picture of the trend of experiences at these ten (twelve) schools. These meetings took place in October 1983, January 1984, May 1984, November 1984 and May 1985.
The November 1984 meeting was unfit to extract experiences from because of its specific nature: it was solely devoted to the problems of testing and examinations. All other meetings gave possibilities to classify the experiences in some way. We have tried to classify the experiences in three classes: when the teachers spoke of difficulties, problems etc. we classified these experiences as '−'.

When the teachers has mixed feelings, or neutral feelings, we classified them as '0'. In the case when teachers were positive or very positive we classified the experiences with '+'.
In this way we get the following table, resulting in the graph below it.

REACTIONS 12 SCHOOLS
(FROM DEC. 82 TILL MAY 85

REACTION # / SCHOOL #	1	2	3	4	5	6	7
1	+	+		+	+		+
2		−		+	+	o	+
3	+	+		+			+
4	−	o			+		+
5	−	o	+				+
6	−	o	+	o	o		+
7	−	o	+	+	+		+
8	+	+		+	+		+
9	o	o		+			+
10	o	o	+	+	+		+
11	+			+	+		+
12				+		+	+
	DEC 82	MARCH 83	MAY 83	OCT 83	JAN 84	MAY 84	MAY 85
	CLASS 4			CLASS 6			EXAM

OR GRAPHICALLY:



When interpreting this data we have to be very careful because of the arbitraries of the classification. But one conclusion is clear: after a introductionary period which caused some doubts about the new program there is a steady increase in the rating of the experiences.

We would like to conclude the experiences of the twelve schools with a quote of a teacher in an article about the changes that took place in his classroom-practice [15]:

"The introduction of the Hewet-*material presses the teacher to compare his own vision on mathematics education with the* Hewet-*view.*
But also the students have a view of mathematics. And this can lead to confrontations, as we experiences. Learning by 'explaining' and 'reproduction' is quite something different from mathematization as a group activity."
"What did we learn from teaching Mathematics A? Quite a bit. Teaching is a dynamic process with unexpected and unforeseen results.
Mathematics education is less self-evident than one assumes. Teaching 'as-always-the-same-rate' is a fable.
Obviously it is possible to lead students to activities which make them experience mathematics as an activity."

For the 52 schools reactions were taken at two conferences in 1985. The question posed to the teachers was a very simple one:

"What are your impressions about the teaching of Mathematics A?"

We have tried to tabulate the results of the latest available sounding in order to get an overview. For this reason we classified the reactions in four classes initially: negative, neutral, moderately positive and positive. However, none of the reactions was negative. This leaves us with three classes: x neutral; xx moderately positive; xxx positive; very positive.
At the same time we took into account the most mentioned critical remarks. In our analysis we found seven points that were mentioned more than once spontaneously.
As our table on the next page shows these points are:
- lack of (basic) skills (6);
- lack of structure (2);
- difficult to extract essentials (3);
- problems with mixed A/B groups;
- too overloaded (5);
  (or, for one school: ample time) (1);
- difficult (7)/too much variety in difficult and easy parts (3);
- problems how to test (7).

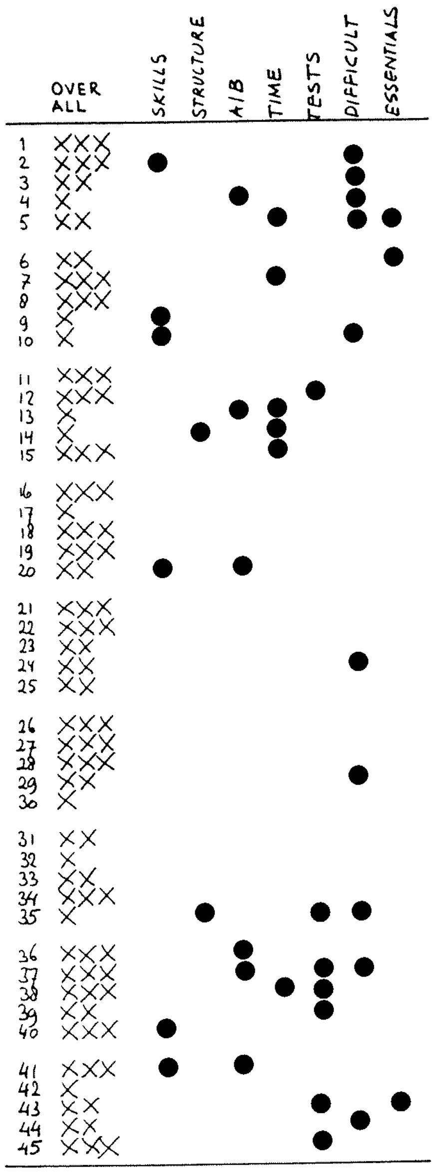
Looking at the totals we notice the following:

x	Neutral:	10	or	22%
xx	Moderately, Positive:	13	or	29%
xxx	Positive, Very Positive:	22	or	59%

based on 90% of schools

The table may lead to the conclusion that the general impression of the experiences is better than moderately positive. At the same time though we should bear in mind that all schools volunteered for these experiments, and are thus by no means representing the total population of 450 schools that introduced the new program in the summer of 1985.
We would like to conclude with a reaction from a student, one year after leaving school:

"It was very stimulating to have the idea to discover, to invent by yourself.
And it came as a big surprise to me that I still know all my mathematics at this date!" [16]



Achievement testing

One of the problems often mentioned by the teachers is constructing achievement tests.
A question asked: *"How do we operationalise the goals of mathematics A?"* This was a hard to answer question while the goals were not completely clear at the moment.
The *Hewet*-report is not very clear on the goals and the first time the goals, were made explicit was in 1985 [17] Nevertheless some teachers had the feeling that mathematization was an activity to be tested, and that the processes were more important than the product, at least in certain cases. In relation to this we once more mention the 'rat-problem'.
And how about reflexion?
Questions like these arose, but the answers were not simple.

Restricted-time-written-tests at the twelve schools

During the first two years of the experiment for the twelve schools more than one hundred tests were collected, including results, problems, and other remarks. Two questions – which are related – were the background for this investigation:
- How well are the goals of mathematics A operationalised in restricted-time-written-tests?
  And of these goals are insufficiently operationalised is this because of:
  - the goals of math A were not clear for the teachers;
  - the goals are not properly operationisable by means of time-written-tests?
- Are the exercises of the tests formulated within the context of the studentmaterials?
  Do they resemble the exercises in these materials, or are new contexts and new problems offered to find out how well the transfer to other problems has been mastered?

The last question can be answered easier than the first one if we restrict ourselves to the collected tests. These tests were made with hardly any help from the *Hewet*-team or from the teachers of the two schools. Of course the teachers had seen the tests from the two schools but it was agreed that no further cooperation should take place in order to get a proper impression of how the different schools constructed their own restricted-time-written-tests.

Of the more than one hundred tests that were collected we will analyse in some detail the 78 problems from 23 tests on the subject '*Matrices*' which represents the overall picture of the collected tests.
In order to answer our second question we have to give some classification that takes into account 'the distance' of the exercises in the tests to the exercises of the studentmaterial.
We have choosen for a classification in the following way:

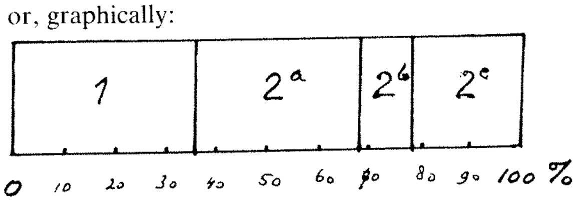
Class 1
Exercises without context; or with hardly any context.

Class 2
Exercises with substantial use of context:
a. Exercises that have strong resemblance with exercises from the booklet.
b. Exercises that have resemblance with exercises from the booklet, but not very obvious.
c. Exercises that have no resemblance with exercises from the booklet.

The distribution of the exercises over the different classes is as follows:

class	exercises
1	28
2a	25
2b	9
2c	16

or, graphically:



Our conclusion, following from analysing the tests as executed by the twelve schools is that roughly 80% of the exercises looks more or less alike the exercises in the booklets. The class 2c exercises that try to test higher goals of mathematics A are represented by 20%.

This leads also to an indication for an answer – as we indicated before – on the first question:

In the restricted-time-written-tests the goals of mathematics A are operationalised in a limited way.

The reasons for this fact are twofold as discussion with the teachers prove:

In the first place the goals of mathematics are not completely clear to teachers and, even more important:

It is very difficult to construct tests for mathematics A, especially because of the fact that some goals cannot be operationalised properly by means of restricted-time-written-tests.

This last observation leads us to our explorative study to find out other forms of assessments tasks that take into account not only those aspects that can be examined by means of restricted-time-written-tests but also of those aspects which need to be assessed in some other way.

Principles of tasks

As soon as the *Hewet*-team became involved in constructing achievement tests – in October 1981 in the two-schools-phase – the tension between the goals of mathematics A and the restricted-time-written-tests became clear. It was considered normal to accept this practice in Dutch schools not to disturb the daily routine more than necessary.

From our analysis of tests of the twelve schools it becomes clear that the testing tends to emphasize the 'lower' behaviour levels, such as computation and comprehension. This is not a specific Mathematics-A problem. As Wilson states: Mathematics teachers often state their goals of instruction to include all cognitive levels. They want their students to be able to solve problems creatively. But too much of their testing consists only of recall of definition, facts, and symbolism. [18]

But in Mathematics-A the problem is even more serious because of the nature of the goals.

Mathematization, Reflexion, Inventiveness and Creativity are activities that are essential for mathematics-A and very hard to test in the restricted-time-written-test.

The teachers in the experimental schools were hindered by several facts. In the first place the goals of mathematics were not clearly stated, which made teachers rely heavily on the textbookmaterials. In the second place no good exercises outside the booklets were available and only few teachers created 'new' exercises as we noticed in the preceding chapter. In the third place the teachers were under heavy time pressure as there were hardly any provisions for teachers participating in the experiment: preparations for teaching mathematics-A took much time compared to the old program. In the fourth place the tradition of testing mathematical skills in Dutch schools didn't inspire teachers to get involved in tests that are different from the ones in the textbooks.

During the experiment the goals of the teaching of mathematics-A became clearer. As did the problem of assessment testing. The words of the Cockcroft-report [19] that assessment (at 16^{+}) is to reflect as many aspects as possible, it needs to take account those aspects which need to be assessed in another way than restricted-time-written-tests got a very concrete meaning for the *Hewet*-team and some teachers as well. The team tried to counter the problems with aid of the teachers in the following way:

In the first place an attempt was taken to clearify the goals of mathematics A.

Secondly, and of more direct help to the teachers, exercises were sampled and distributed to the field in order to present teachers 'new' exercises.

The principles behind the development of alternative tasks were the following:

1. The tests should improve learning.
2. The tests should be such that the candidates show what they know (positive testing).
3. The tests should operationalise the goals of the mathematics A-curriculum
4. The quality of the test is not in the first place measured by the possibility to score in an objective way.
5. The tests should fit in the usual school-practice.

The last principle is of a pure pragmatic nature, in order to enlarge the possibility that teachers will introduce these tasks in their school-practice.

The fourth principle has to do that we do not accept all restrictions to testing when the objectivity seems to be the major goal; actually tests that can be scored in an objective way do in general not obey our first three principles. These principles hold – in our opinion – in general but are essential when we relate them to our methodological choice for the experimental material for mathematics A.

The third characteristic point of a realistic curriculum was the large contribution from students to the course by their (mental) productions and constructions. Constructing, reflecting, anticipating and integrating are fundamental functions of the students own production.

So also in tests and tasks there should be a possibility for students to construct and produce, in a positive way, in order to improve learning.

There is little doubt that developing tests is a complex matter if we want to stick to our five principles. But at least we should try to stay as close as possible to our principles.

Two-stage-test

This test tries to combine the advantages of the traditional restricted-time-written-test and the five basic principles as pointed out before.

The characteristics of the restricted-time-written-test are:

- The test stresses 'what you don't know'.
- At a fixed time the same test for all students at school.
- All students have to make the test within a fixed time limit.
- Most of the time a lot of attention is given to the lower goals.
- Questions are of open-structure: short- and long-answer questions.
- Scores are considered relating objective (for mathematics).

These are also the characteristics of our first stage of the two-stage-test. From these and from our basic principles the characteristics for the second-stage of our test follow:

- The test is done at any convenient time at home.
- The amount of time on the test is almost unlimited.
- The test stresses 'what you do know'.
- Much attention is given to the 'higher' goals: interpretation, reflexion.
- Questions are of open-structure: long-answer questions and essay-type-questions.
- Scoring is considered as difficult and not objective at all.

The two-stage-test, inspired by an idea of Van der Blij, is one test, with open questions and essay-questions. The first stage is executed like a traditional restricted-time-written-test. The students are expected to answer as many questions as possible a set time, and within a fixed time limit, being 45 minutes.
In fact, the student has complete freedom in making the choice of questions he wants to tackle during this time. But the first half of the test has mainly open questions, and the second half also essay-questions so one may expect the students to answer the open questions in the first place.

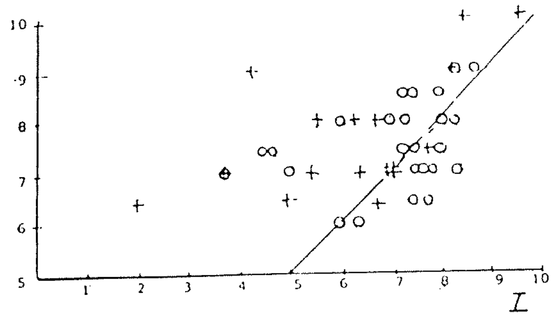
After this first stage is finished the teacher takes the tests home to score them. Than the tests are handed back to the students who know their score, and their biggest mistakes (only).
Now the second stage takes place: the students take the test and their answers back home and get the opportunity to make the test again in a way that they prefer. Basically they have complete freedom in the way they want to answer the questions: question after question or an essay with implicit answer or any combination.
After a certain time – around three weeks – the students have to turn in their work and a second scoring takes place. In this way this test delivers the teacher (and the students) two grades: one for the first stage, and one for the second stage.

About objective scoring

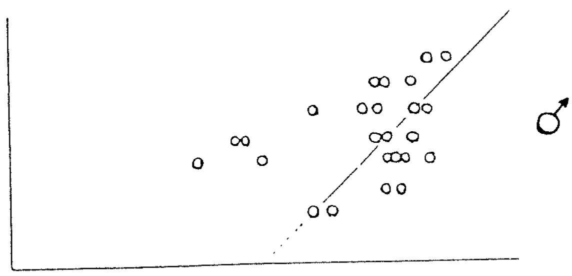
Scoring the first stage of the test was no problem for the teacher as this was an ordinary restricted time written test. And there is a general feeling that these kind of tests are – at least in mathematics – fairly objective scorable.

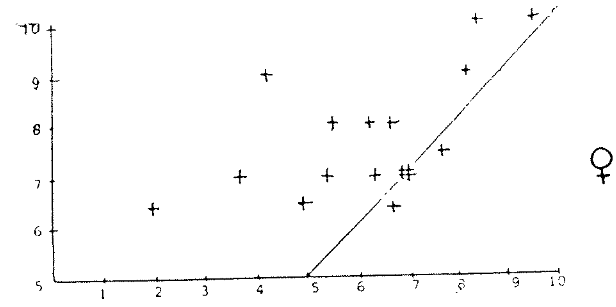
Scoring the second stage is a problem. The teacher decided to read the work of the student and to make positive and negative notes, resulting in a score. So no precise table, just an impression that could be defended by pointing to these positive and negative points.

Comparing the scores of the students, there are some remarks to be made. This remarks are not based on the experiences with this test and this forty students only, but follow from other alternative tests as well. First let us look at the graph showing the total results:



It shows clearly that the second round gives higher results, according to what we expected. When we make seperate graphs for boys and girls we get:





From these graphs one would like to conclude that girls used the possibilities of the second stage better than boys.
The results of the girls in the first stage were considerably lower than those of the boys; but not so in the second stage.
The best results were actually from girls.

The questions remains how objective are the scores of the second stage. Here we have open questions, essay-type answers, and we may expect problems in scoring these questions.
We send out five of these tasks to twelve teachers, spread all over the country. These teachers had no knowledge at all about:

- the score for the first stage;
- the scores in general of those 5 students;
- how to score the task.

The only question asked to the teacher:
Score this task.
The reactions showed that most teachers felt very uncertain. Not only about the mathematics in the task, but also the complete lack of references for scoring was a big problem.
Some teachers decided just to read the tasks carefully and give intuitively a score. Others made a complicated and precise model for how to score.
The results were:

Student	Teacher 1	2	3	4	5	6	7	8	9	10	11	12
1	6	5	$6\frac{1}{2}$	7	6-	7	6	7	$6\frac{1}{2}$	8	7	6,5
2	7	8	7	8	6	7+	8	6	7	$7\frac{1}{2}$	6	7,25
3	8	6	7	8	$6\frac{1}{2}$	8-	8	$8\frac{1}{2}$	$7\frac{1}{2}$	$8\frac{1}{2}$	6	7,5
4	9	8	9	9	$7\frac{1}{2}$	8	8	9	$8\frac{1}{2}$	8+	8	8,25
5	10	$8\frac{1}{2}$	8,75	9	8	8+	9	9	10	$9\frac{1}{2}$	9	8,75

Of course it is very difficult to draw any conclusions from this table. There seems to be more agreement. Look for instance at student 4:
9 8 9 9 $7\frac{1}{2}$ 8 8 9 $8\frac{1}{2}$ 8+ 8 8,25
but there are differences as well. On first sight the scores of student 1 seem to differ quite a bit, but if we take away the two extremes there is wide agreement:
6 $6\frac{1}{2}$ 7 6⁻ 7 6 7 $6\frac{1}{2}$ 7 $6\frac{1}{2}$.
Because of the fact that there was no answer-model, or any other reference that could support the teachers the spread of scores doesnot seem very wide. If we assume that such an answer-model would give more or less the same average and spread we figured out what scores we would have if:

- each school has the same average (7.7);
- each school has the same S.D. (1,0).

This lead to the following table:

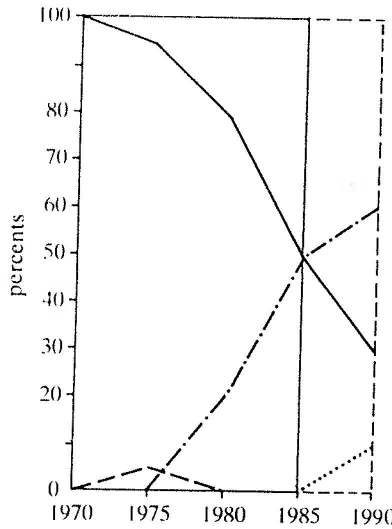
Student	Teacher 1	2	3	4	5	6	7	8	9	10	11	12
1	7,5	7,2	6,6	6,1	6,3	6,2	6,5	6,7	6,2	5,9	7,0	6,2
2	6,7	6,4	7,0	7,4	7,0	8,4	7,0	6,7	6,7	7,9	6,1	7,2
3	6,7	7,9	7,4	7,4	7,7	6,9	7,0	7,3	7,8	7,9	8,2	7,5
4	8,4	7,5	8,2	8,8	8,4	8,4	9,0	8,2	8,3	7,9	8,6	8,3
5	9,2	9,4	9,4	8,8	9,1	8,7	8,7	8,6	8,9	8,9	8,6	9,1

Only at a few places the *order* is not in complete accordance (looking vertically).
If we look horizontally we see that the differences between the scores of one student is not too big. Actually, student nr. 5 should have a 9 with all teachers if the scores were rounded off.

The results should be studied more carefully, and the ways the teachers made up their score as well.
Nevertheless one is tempted to conclude that the fear for lack of objectivity when scoring more open student-tasks seems not completely justified. Especially when we take into account the content-related advantages, and the fact that most other restricted-time-written-tests will be scored in an objectve way, there seems to be no valid arguments to oppose such tasks other than time-consuming.

Towards realistic mathematics education

The *Hewet*-project seems a step forward to more meaningful mathematics education. The intuitive exploration leading to new mathematical concepts, the need to shift the focus of attention from product to process, the crucial role of the discussion between students, and the productive role of the students, they all make that mathematics-A breaks away from routine exercises and non-purposeful tasks. But there are other strong movements towards more realistic mathematics education. De Jong [20] published in 1986 his research about the effects of the Wiskobas-group. And his conclusions are worth some attention. Realistic methods introduced for the first time in 1975, now take 50% of the market as is shown by the following graph:



——— mechanistische methoden
—·—·— realistische methoden
– – – – structuralistische methoden
········· hybridische methoden

At the same time there is a strong decline in mechanistic methods. Remarkable also is the fact that the structuralistic approach [Bourbaki; Papy] never had any chance in the Netherlands. Largely because of the Wiskobas-activities. But there is more to come. We mentioned already the National Plan wherein we find a strong support for the ideas that we call 'realistic'. And even more recently, the Scientific Council for Advise to the Government made clear that also that part of math-education that has been neglected so far – the 12–16 year age group – will face major changes

in the near future [21]. And again these changes will be in the direction that was indicated by Wiskobas and the *Hewet*-project.
But developmental research combined with the development of student texts is not sufficient. Support is needed to offer in-service-training for those who are already active as teachers. Support is needed to introduce students and teachers to the use of computers and calculators in the classroom.
And the teacher-training for primary school teachers is a major point of concern. But that was already clear from our first example.

OW & OC '5 years'. A small group of IOWO-survivors is taking its part, and a responsible part, in the development towards realistic mathematics education. And with the cooperation of many others, including teachers, teacher-trainers, curriculum developers, researchers, computer experts and most important pupils, we are convinced that the changes are a change in the proper direction.

Literatuur

[1] Streefland, L.: *How to teach fractions so as to be useful*, OW & OC, Utrecht University, The Netherlands, 1984.
[2] Brink, F.J. van den and L. Streefland: *Young children (6-8) – Ratio and proportion,* Educational Studies in Mathematics 10, p. 403-420, 1979.
[3] Treffers, A. and E. de Moor: *10 voor de basisvorming*, OW & OC, Utrecht University, The Netherlands, 1984.
[4] Lange, J. de and M. Kindt: *The Hewet-project – Report on an experiment to a new curriculum for pre-university students*, Zentralblatt für Didaktik der Mathematik 16 (2), p. 74-79, 1984.
[5] Lesh, R., M. Landau and E. Hamilton: *Conceptual Models in Applied Mathematical Problem Solving* in Lesh, R. and M. Landau (eds): *Acquisition of Mathematics Concepts and Processes*, Academic Press, New York, 1983.
[6] 't Hart, M.: *Ratten*, A.P. Querido, Amsterdam.
[7] Lange, J. de: *1808*, Nieuwe Wiskrant, 3e jrg. nr. 2, OW & OC, Utrecht University, The Netherlands, 1983.
[8] Treffers, A.: *Three dimensions*, Reidel, Dordrecht, 1986.
[9] Treffers, A. and F. Goffree: *Rational Analysis of Realistic Mathematics Education.* In Streefland, L. (ed), *Proceedings of the ninth International Conference for the Psychology of Mathematics education,* OW & OC, Utrecht University, The Netherlands, 1985.
[10] Hilton, P.: *Current Trends in Mathematics and Future Trends in Mathematics Education*, For the learning of Mathematics 4 (1), p. 2-9, 1984.
[11] Courant, R., H. Coxeter etc.: *On the mathematics curriculum of the high school*, The Mathematics Teacher (3), American Mathematical Monthly (3), 1962.
[12] Fehr, H.: *New Thinking in Schoolmathematics*, Organisation for European economic cooperation, Paris, 1961.
[13] Kindt, M. and J. de Lange: *Realistic Math for almost all.* In Lange, J. de (ed) *Mathematics for all ... in the computerage, Proceedings of C.I.E.A.E.M. 37*, OW & OC, Utrecht University, The Netherlands, 1985.
[14] These reports by H. Verhage were made during the in-service-teacher-training-courses during the Hewet-project.
[15] Koerts, E.M. and M.C. Visser: *Betekenisvol leren met Hewet*, Nieuwe Wiskrant jrg. 3 nr. 3, OW & OC, Utrecht University, The Netherlands, 1984.
Taken from a survey carried out by H. van der Kooij (not published).
[17] Lange, J. de: *Toetsen bij Wiskunde A*, Nieuwe Wiskrant jrg. 4 nr. 4, OW & OC, Utrecht University, The Netherlands, 1985.
[18] Bloom, B., T. Madaus and G. Hastings: *Handbook on formative and summative evaluation of student learning*, McGraw-Hill, New York, 1971.
[19] Cockcroft, W.H.: *Mathematics Counts, Report of the Committee of Inquiry into the Teaching of Mathematics in Schools*, H.M.S.O., London, 1982.
[20] Jong, R. de: *Wiskobas in methoden*, OW & OC, Utrecht University, The Netherlands, 1986.
[21] Blij, F. van der and A. Treffers: *Basisvorming in het onderwijs: Rekenen-Wiskunde*, Den Haag, The Netherlands, 1985.