Gender differences in mathematics achievements in Dutch primary schools

On the search for features of mathematics education that are important for girls

Marja van den Heuvel-Panhuizen
Freudenthal Institute, University Utrecht

“We know others’ work as published product; we know our own better as work-in-process”
(Glaser and Strauss, 1967, p. 9)

1 Introduction

Twenty five years ago, in The Netherlands the first steps were taken towards the reform of mathematics education, which later became known as “Realistic Mathematics Education” (RME). Characteristic of this new approach to mathematics education is the rejection of the mechanistic, procedure-focused way of teaching in which the learning content is automatized in meaningless small parts and where the students are offered fixed solving procedures to be trained by exercises, often to be done individually. RME, on the contrary, has a more complex and meaningful conceptualization of teaching. The students, instead of being the receivers of ready-made mathematics, are considered as active participants in the teaching-learning process, in which they develop mathematical tools and insights. The basis for this new approach to mathematics education emerged from Freudenthal’s (1971, 1973) idea of mathematics as a human activity which he connected with the principle of guided reinvention. This means that, in RME, the own constructions and productions of students play a central role. As Treffers (1987) indicated in his description of the theoretical framework of RME, the contributions of the students are one of the five elements that constitute RME curricula. The others elements are the major place of contextual problems and real-life situations by means of which the students can both constitute and apply mathematical concepts, the use of models by which the gap between the informal, context-connected mathematics and the formal mathematics can be bridged, the interactive character of the teaching process, and, finally, the intertwinement of various learning strands.

Although RME has left already its mark upon today’s primary school mathematics education in The Netherlands — more than three-quarters of the Dutch primary schools now use a mathematics textbook that was inspired to a greater or lesser degree by this reform movement — RME is still under development. One is still working on the implementation of RME in classroom practice and the revision of textbook series. The process of theory development is also not yet finished. Continuously, new adjustments are made of the realistic domain-specific teaching/learning theory and blind spots in this theory are detected and filled in. An example of such a blind spot is mathematics education to girls. Besides the more general characteristic of RME that the problems should be imaginable for students and should always fit within the child’s world,

the aspect of gender has never been a special topic of investigation within RME at primary school level. It were the results of two successive National Assessments of the Educational Achievements (PPON) conducted by the National Institute of Educational Measurement (CITO), which, in The Netherlands, placed the gender issue on the research agenda. The results from these gauges arose the presumption that the girls maybe could have less profit of the mathematics education reform movement in primary schools than the boys. To investigate this and to find out what would be powerful learning environments for girls — which possibly could result in adjustments of the RME theory — the MOOJ Project has set up. The project is funded by the Dutch Ministry of Education and conducted by the Freudenthal Institute of Utrecht University and the Center for the Study of Education and Instruction of the State University of Leiden. The CITO is also involved in the study.2

The complete project will contain three stages:

I A further exploration of the gender differences in mathematics achievements at primary school level

II A more thorough investigation at a number of schools

III The set up of a teacher enhancement project based on the results of stage I and II.

Making this study supportive to the improvement of classroom practice, is the main founding principle of the MOOJ research.

Present article will provide a description of stage I. It will start with some background information about how the study came into being. Then, much attention will be paid to the aim and the research design of stage I and its results. Finally, after placing these results in an international perspective, the suitability of RME for girls is questioned.

2 The research questions and set up of stage I of the MOOJ Project

A short sketch of what preceded to the MOOJ Project

The actual reason why in 1995 the MOOJ Project was started were the results of the second PPON conducted in 1992 (Bokhove, Van der Schoot and Eggen, 19963). This PPON was the second national assessment that had shown that in Dutch primary schools the girls have lower average scores for mathematics than the boys. The same was revealed by the first PPON held in 1987 (Wijstra (ed.), 1988). Then, Freudenthal, inspired among others by the Dutch working group “Women and mathematics”, was the main advocate who pleaded for further research on this gender difference. According to him this further research should especially focus on the composition of the research group. In the PPON studies, namely, were only students involved from regular primary schools and not from schools for children with special needs. The problem with this restricted focus is that in The Netherlands more boys are placed in special needs schools than girls. A consequence of this could be that the composition of the research group might be in the disadvantage of girls. However, it never came to a concrete plan to test this “special needs-hypothesis”. The thing that was done, was that in

2 The Utrecht part of the project team consist of Marja van den Heuvel-Panhuizen, Koeno Gravemeijer, Adri Treffers and Leen Streefland and the Leiden team consist of Harriet Vermeer, Monique Boekaerts and Gerard Seegers. The CITO staff members that are involved in the study are Henny Uiterwijk and Joop Bokhove.

3 The results of this second PPON were presented at the PANAMA Conference in 1994. This is a yearly Dutch conference on primary school mathematics education.
1992 a separate PPON for schools for children with special needs was held. The results of which (Kraemer, Van der Schoot and Veldhuijzen, 1996) made clear that also in special needs schools the boys’ scores in mathematics surpassed the girls’ scores. Although this outcome can exclude some possibilities regarding the gender differences in the complete population, it cannot give enough support to reject the “special needs-hypothesis”. In other words, the question has not been answered yet, whether the differences in mathematics achievements between girls and boys, found in the regular PPON studies, are a result of the composition of the research group, in combination with the gender differences in the placements in special needs school.

Aim of the MOOJ project and especially the aim of stage I of that research project

This question will also not be answered in the MOOJ research. In the research proposal (Boekaerts, Bokhove, Gravemeijer and Treffers, 1995) that was written in consequence of the second PPON and that was submitted by the Dutch Ministry of Education, it was actually taken for granted that at primary school level girls are not that good at mathematics as boys. All the previous national assessments came namely to this result. Therefore, the proposed study would be especially aimed at the search for possibilities to get rid of, or minimize this inequality. The aim of the complete MOOJ Project is: to trace mechanisms that are responsible for the creation of this inequality and find footholds for the improvement of the mathematics scores of girls. By means of a thorough investigation of the teaching methods it will be examined which teaching, school and teacher factors, in collaboration with student characteristics, can cause gender differences in mathematics achievements. Classroom observations will form the heart of this research. For these observations schools will be selected in which the girls do relatively well. This means schools of which the average mathematics score of the girls, at least, equals the one of the boys. In the MOOJ Project these schools are called “girls schools”. These schools will be compared with schools in which the boys’ average mathematics score is higher than the one of the girls. These school are called “boys schools”. Because of the fact that the existence of both types of schools was a requirement for carrying out the research, the first question to be answered was whether these schools exist or not. This was especially true of the girls schools. The goal of stage I of the research was to investigate this. At the same time the purpose of stage I was to provide an overview of the gender differences in mathematics achievements at the end of the primary school.

Research design stage I

In order to investigate the differences in mathematics achievements a continued analysis was carried out of the CITO data base from the CITO final test for primary school. This CITO test is meant for providing an individual student score for making a decision about the enrollment in secondary education. The test is not compulsory, but is administered in approximately 70% of the Dutch schools. This means that yearly about 5000 schools use the test. In other words, each year, a little bit more than 100,000 grade-

---

4 The name of the research project was not used in the research proposal and was chosen not until the end of stage I. The letter M in acronym “MOOJ” stands for “Meisjes” which means “Girls” and the letter J stands for “Jongens” which means “Boys”. The two letters O are used for making the two gender signs.

5 The Utrecht part of the research team had the responsibility for the execution of this stage of the research. The complete research report of stage I can be found in Van den Heuvel-Panhuizen (1996).
six students do this test. Along with items on mother language and information processing, the test contains 60 items on mathematics, divided in three parts. The test items are presented in a multiple choice format.

For the MOOJ Project, the 1993 through 1995 mathematics data of the CITO final test for primary school were analyzed on three different levels: (i) the individual student level, (ii) the level of the discrete test items, and (iii) the aggregated school level. The first two analyses were meant for getting insight in how the mathematics achievements differ between girls and boys. The third analysis was aimed at the selection of schools for stage II of the MOOJ Project.

(i) The analysis on individual student level included that for each year, for both sexes separately, the average percentage correct answers (p-value) and the standard deviation was calculated.

(ii) The extra analysis on test item level was done because of the fact that the subscores of the CITO final test for primary school are rather rough to reveal specific abilities of girls. This analysis included the calculation of the p-value of each test item for both sexes separately, followed by the calculation of the difference between both the p-values. Then, for all the three research years, those items have been selected of which the difference the largest or the smallest was, or possible in the advantage of the girls. Next, a qualitative analysis was carried out on these “extreme” items which was focused on the mathematical content of the problems. For this analysis no particular criteria were formulated in advance. It was a very open analysis in which the following procedure was followed: (a) repeated studying of the problems (reading and re-reading the problems and imaging oneself again and again what the problem is asking of the student when he or she is doing the problem) until particular characteristics emerge, and (b) then trying to identify these characteristics in other problems, (c) possibly followed by a revision of the characteristics.

(iii) The analysis on school level included the calculation of the average score of the girls and the boys for each school. The t-test was used to determine whether both scores differed significantly. The magnitude of the t-value was used to distinguish several categories of schools, running from schools were the boys did better than the girls, to schools were the scores were approximately equal, and finally, to schools were the girls’ score was a little bit better than the boys’ score.

3 The results from stage I of the MOOJ Project

The results from the analysis on student level

The results that were found in the analysis on student level confirmed the findings of the two PPON studies. In each year under investigation the grade-six boys had better mathematics scores on the CITO final test for primary school than the grade-six girls (see Table 1). In 1993, the boys answered 71% of the sixty problems correctly, while the girls came up with 65% correct answers. This 6% difference in the total score for mathematics is a little bit more than a quarter of the standard deviation of the scores. The same results were found for the total scores in the other years and for the five subscores that CITO distinguishes for mathematics. These subscores include: basic knowledge of numbers; mental arithmetic; operations; fractions, percentages, and ratios; measurement, time, and money. The gender
differences between the subscores were rather small. Probably this was caused by the way in which the problems for the different mathematics domains have been operationalized in the test.

Regarding to the gender differences in the subscores, the PPON studies were more revealing. These studies showed remarkable differences between the boys’ subscores and the girls’ subscores, which, however, is not surprisingly. In the PPON studies the measurement of the sub-abilities is much more precise than in the CITO final test for primary school. In the two PPON studies respectively 27 and 29 different scales were used. The scales on which the girls performed better than the boys were less in number, but at least there were some. Moreover, it did not look as a matter of chance, because the pattern was the same in both of the studies. In the final assessment in grade six of the first PPON in 1987 (see Wijnstra (ed.), 1988) on 21 scales (out of 27) the scores of the boys and the girls differed significantly (p < .05) in the advantage of the boys. The largest differences were found in the scales basic knowledge and understanding of measurement, applications of measurement, basic knowledge and understanding of ratio, applications of ratio, estimations, and basic knowledge and understanding of whole numbers. The only scale on which the average score of the girls was significantly higher than the one of the boys was the scale written algorithms for division. On the three other scales for written algorithms the differences were not significant, but in two of them the difference was in the advantage of the girls. The results on the second PPON (Bokhove, Van der Schoot, and Eggen, 1996), held in 1992, showed that there still exist large gender differences in mathematics achievements. In this second study, all the differences found in grade six were significantly in the advantage of the boys. Only in four of the five scales for written algorithms the average scores of the girls were higher than the ones of the boys, although not significantly. Characteristic is, however, that on the fifth scale for written algorithms, applications of written algorithms, the average score of the boys was significantly higher again. Regarding to the nature of the differences, the PPON results resemble the findings from other studies.

Table 1: Mathematics achievements of boys and girls on the CITO final test for primary school

<table>
<thead>
<tr>
<th>Cito final test for primary school</th>
<th>1993</th>
<th>1994</th>
<th>1995</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>girls</td>
<td>boys</td>
<td>girls</td>
</tr>
<tr>
<td>Total score mathematics (60 test items)</td>
<td>n=50111</td>
<td>n=49411</td>
<td>n=52600</td>
</tr>
<tr>
<td>Basic knowledge numbers</td>
<td>67</td>
<td>74</td>
<td>63</td>
</tr>
<tr>
<td>Mental arithmetic</td>
<td>66</td>
<td>72</td>
<td>67</td>
</tr>
<tr>
<td>Operations</td>
<td>67</td>
<td>74</td>
<td>63</td>
</tr>
<tr>
<td>Fractions, percentages, ratios</td>
<td>62</td>
<td>69</td>
<td>72</td>
</tr>
<tr>
<td>Measurement, time, money</td>
<td>59</td>
<td>67</td>
<td>60</td>
</tr>
</tbody>
</table>

*The scores in this table only involve the students of which the sexe was filled in on the test page. In each year this information was lacking for about 2% of the students.*

Regarding to the nature of the differences, the PPON results resemble the findings from other studies

---

6 Each time when in the following a difference is called significant, then this means p < .05.
(see, for instance, Marshall, 1984), that girls do better on bare problems and boys have better scores on application problems. Because of the fact that the PPON scales for written algorithms also half consist of simple context problems, this conclusion could not be drawn by the PPON researchers. Nevertheless, the results gave them support to raise the conjecture that the gender difference in mathematics achievements more has to do with a difference that also emerges in language achievements: girls score higher on tasks which includes “executive aspects of language”, but have weaker scores on tasks which ask for “the ability to manipulate verbal concepts” (see Wijnstra (Ed.), 1988, p. 158).

The results from the analysis on item level

The extra analysis that was carried out on item level revealed that the differences in p-values varied from +.26 to –.04.\textsuperscript{7} The two groups of extreme items\textsuperscript{8} that have been determined based on the differences in p-value can be distinguished in “boys problems” which were better done by the boys than by the girls, and “girls problems” in which the p-values of boys and girls were about equal, or a little bit higher for the girls.

![Figure 1](image_url)

**Figure 1:**
- Left part: Example of a “boys problem” from the CITO final test for primary school (year 1995, part 3, item 3; difference in p-value .26)
- Right part: Example of a “girls problem” from the CITO final test for primary school

---

\textsuperscript{7} The difference was calculated by taking the p-value of the boys minus the p-value of the girls.

\textsuperscript{8} These extreme items were found by selecting those items of which the gender difference was the largest. This was done in each of the three parts of which the CITO final test for primary school consist of yearly. For each part of the test approximately four items were chosen. In order to get for both sexes about the same number of items, for the girls were also items selected of which the difference was nearby zero. Finally, this selection resulted in a collection of 34 problems that were better done by the boys and 36 problems that were done relatively well by the girls.
For each category an example is shown in Figure 1. The left part of this figure contains an example of a boys problem. In this problem the students have to determine the distance between a certain kilometer peg where a car has stopped and a telephone which is at a different kilometer peg. The problem has to be answered in meters, whereas on the pegs kilometers are used with a notation in decimals. This problem was extremely better done by the boys. The difference in p-value ran up to .26.

The right part of Figure 1 contains an example of a girls problem. Given in this problem is what a person who likes to buy a camera has saved in four subsequential months. The students have to calculate what is the shortage for buying the camera. The difference in p-value of this problem was .04 in the advantage of the girls.

The analysis of the extreme items made clear that the boys perform better than girls on
- problems which ask for daily-life knowledge on numbers and measures
- problems in which large numbers with many zeros are used
- problems in which different numbers or different units of measurement are used
- problems which have possibilities for “tinkering” with numbers
- problems which ask for reasoning backwards

The girls perform equally well as the boys or a little bit better than the boys on
- problems which ask for accuracy
- problems of which the text is complex
- problems which ask for (reflection on) strategies and not for calculations
- well-known problems which refer to standard procedures
- straight-away problems
- problems which refer to shopping situations.

Not only could problem characteristics clearly be distinguished into boys characteristics and girls characteristics, the problems’ pattern of scores appeared also to be very stable over the years under investigation. An example of this can be found in the problems about meters, like, for instance, gas meters, in which the meter heights have to be compared. Most of these problems ask for precise calculations. This characteristic is recognizable in the pattern of scores. Each year the girls performed slightly better on these problems than the boys did.

These results from the analysis on item level can be considered as important indications for how mathematics education can be optimized for girls. As a matter of fact, these findings formed important building blocks for the set up of the research design of stage II of the MOOJ Project.

---

9 By “tinkering” is meant the application of smart strategies. This tinkering mostly implies the application of another strategy than the one that is overtly presented in the problem. Take for instance a problem like “7980 x 0.25 is approximately ....”. Solving the problem by tinkering would mean dividing by four instead of multiplying by 0.25.

10 See CITO final test for primary school, year 1993, part 3, item 5; CITO final test for primary school, year 1994, part 1, item 4; CITO final test for primary school, year 1995, part 3, item 19.

11 The average p-value of these problems were: 1993: j=.78, m=.97; 1994: j=.75, m=.78; 1995: j=.73, m=.75.
The results from the analysis on school level

In the analysis on school level, seven categories of schools were distinguished based on their t-value of the difference between the mathematics scores of the boys and the girls. The criteria used for establishing the categories are roughly connected with the levels of significance that belong to a certain t-value and a certain size of the school population.\textsuperscript{12} Before for each school both the average scores were calculated, first the schools were removed from the data base which did not fill in the sex of the students and the schools with 10 or less students. Mostly these are starting schools with only a few students in grade six, or Dutch schools in other countries.

Table 2 shows for the years 1993 through 1995 the distribution of the categories over the schools. It turned out that it was not an equal distribution. Only in 1 or 2% of the schools the girls did rather better or clearly better than the boys, whereas in a quarter of the schools this was the case for the boys. Globally spoken, there were two asymmetric halves: in half of the schools the boys were better and in the other half the scores were approximately the same.

Since extreme scores have a great influence on the average score, the same analysis was also carried out minus the 10% strongest and the 10% weakest students. This procedure is done two times: with the removal of the extreme students on student level and the removal of the extreme students on school level. Another reason for excluding these students was that the focus within the MOOJ research will not be on extreme students.

The results of the analysis with the reduction on student level made clear in an indirect way that the group of 10% weakest students contained more girls than boys. The leaving out of the 10% strongest and 10% weakest students resulted namely in average decrease of 50% of the number of schools in which the boys had rather better or clearly better scores than the girls. At the same time, however, this resulted to an average increase of almost 50% of the number of schools in which the girls did better than the boys. Regarding to the reduction on school level the decrease of the boys schools was about the same as for the reduction on

\textsuperscript{12} As a starting point for this was taken an average group of 27 students. If the t-value is over 1.7, then (in the case of two-tailed testing) the level of significance is < .10 and if t-value is over 2, then the level of significance is < .05.
student level. For the girls schools the situation with this reduction was different. Compared to numbers of schools without reduction the average increase in the three years under investigation was about 30%.

Another focus of interest related to this analysis on school level was the question whether the gender differences were systematically or not. Did they, for instance, maintain more than one year? Table 3 makes clear for how many schools this was the case in the total group and in the two reduced groups. In spite of the large number of schools that administered the CITO final test for primary school, there was only a small number of schools which had the same pattern over three years. One must take into account, however, that about 40% or 50% of the schools were removed in advance. This could be done for different reasons. One reason could be that a school did not administer the CITO final test for primary school each year under investigation. Another reason could be that in any of the three years the number of students was not over 10, or the sex was not filled in of all the students.

Table 3: Number of Dutch schools with a particular pattern in the mathematics scores of girls and boys

<table>
<thead>
<tr>
<th>Total number of schools (involved in this analysis)</th>
<th>Number of schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>girls 3 years better *</td>
<td>0</td>
</tr>
<tr>
<td>girls 2 years better and 1 year equal</td>
<td>9</td>
</tr>
<tr>
<td>girls 1 year better and 2 years equal</td>
<td>97</td>
</tr>
<tr>
<td>girls = boys</td>
<td>326</td>
</tr>
<tr>
<td>boys 1 year better and 2 years equal</td>
<td>679</td>
</tr>
<tr>
<td>boys 2 years better and 1 year equal</td>
<td>609</td>
</tr>
<tr>
<td>boys 3 year better</td>
<td>227</td>
</tr>
<tr>
<td>others</td>
<td>187</td>
</tr>
</tbody>
</table>

Table 3 gives evidence that the pattern was not very stable over the three years. This, however, is not so surprisingly. Research by Blok (1993) did already reveal that the school averages on the CITO final test for primary school can vary over the years. An even more important reason for the present finding could probably be the very rough procedure that was followed for categorizing the schools. Nevertheless, despite of the found degree of instability, there were also schools in which the pattern is rather stable. These schools were used as the source from which the schools for stage II of the MOOJ Project were selected.

A summary of the findings of stage I of the MOOJ Project

Stage I of the MOOJ Project revealed (again) that there are gender differences in mathematics achievements at the end of Dutch primary education. The most important findings, based on the CITO final test for primary school, were:

– boys had higher mathematics scores than girls:
  the average total score of the boys was 6% higher than the average total score of the girls
– there were boys’ schools and girls’ schools:
  in half of the schools the average boys’ score was higher than the average girls’ score, and in the other
half of the schools the average scores of both sexes were about equal; so there were two asymmetric halves
- **there were girls’ problems and boys’ problems:**
  - some of the problems were better done by the boys and some were relatively well done by the girls.

### 4 The Dutch situation from an international perspective

*Gender differences in mathematics achievement*

Although the international reports about gender difference in mathematics achievements are not always unequivocally, there is a lot of agreement regarding to young children. At primary school age mostly no gender differences were found for mathematics (see Leder, 1992). To put it even stronger, female primary-school students are often better at mathematics than boys. In a review of English/American gender studies at primary school level, Kimball (1989) found that only in one of the nineteen studies the boys did significantly better than the girls, and in six studies a significant higher mathematics performance for girls was reported. The same was found in several German studies (see Richter, 1995). In The Netherlands, however, the situation seems to be totally different from the one in other countries. Both the PPON studies as well as the MOOJ study showed that in our country there are already
differences in grades 3 and 6. How different this situation is from other countries was recently confirmed by the results of the Third International Mathematics and Science Study (TIMSS) (see Mullis et al., 1997; Knuver and Doolaard, and Matthijssen, 1997). This was especially true of the results found in grade 4 (see Figure 2). For this grade, The Netherlands, together with Korea, was the top leader of the countries in which the differences in mathematics achievements were significantly in the advantage of the boys. The only positive thing was that the results for grade 3 were different. For this grade no differences were found between the mathematics score of girls and boys in The Netherlands. On the other hand, the difficulty with this result is that only 20% of the items of the international TIMSS is judged to be in alignment with the Dutch curriculum for grade 3. For grade 4 half of the items were judged as suitable.

Gender differences in strategies
In contrast to the number of investigations into the gender differences in performance scores in mathematics at primary school level, less research has been done regarding to gender differences in the application of strategies. About this issue still exist many questions. Nevertheless, there are some very interesting starting points for further research. The study done by Fennema et al. (1990) offers, for instance, some valuable indications. In this research carried out in grade 1, the boys scored significantly higher than the girls did on problems which included several operations or extraneous numbers. Fennema et al. suggest that the reason that boys outperformed girls on non-standard tasks might be that the boys are more autonomous than girls in their learning of mathematics. Since the researchers only had written answers they could not be sure about this conclusion. For drawing such a conclusion more information is needed about the students’ strategies. The strategy study carried out in 1991 by Metz-Göckel et al. which is mentioned by Richter (1995) is revealing in this respect. In this study, video-taped observations gave evidence that girls more often than boys tried to get understanding of the structure and nature of the problems, whereas boys more often demonstrated a trial-and-error strategy and were more focused on a quick solution. Recently, Carr and Jessup (1997) found that first-grade girls more often than boys made use of manipulatives or their fingers. Boys, on the contrary, more frequently solved the problems by using the knowledge they already know by heart. Another striking finding in this study was that the boys were not more aware of their strategy than the girls were. Moreover, it was found that when the students had to do the problems in group work, the girls abandoned the use of manipulatives or their fingers.

The above mentioned gender differences in strategies are in tune with the differences found between the boys problems and girls problems in stage I of the MOOJ Project. Think, for instance, of Fennema et al.’s conclusion that boys are probably better in non-standard problems which need autonomous learning. This corresponds with the finding in the MOOJ Project that boys were better in problems which can easy be solved by some “tinkering”. A requirement for this is, however, that you feel yourself free not to solve the problems in the standard way. Another similarity has to do with the use of already known knowledge and the use of daily-life knowledge. These similarities, and others, found between the above mentioned strategy studies and the outcomes of the analysis on test item level give strong indications that further research is needed at this point.

5 Taking a position as a starting point for continuing the research

Overlooking the results of stage I, the need was felt for some reflection and taking a position before continuing the research. The question that arose was: how should one see these results? To what degree gender differences in mathematics achievements can be accepted? Aside from the doubts which one can have about the correctness of the results — a different assessment tool with different problems could have lead to a different result — this question is difficult to answer. For instance, how should be decided whether a 6% difference is acceptable or not for a particular school subject. However, if this question is asked in a different way, then it probably can be answered more easily. Important is namely to realize that in any case it is unacceptable that by a certain approach to mathematics education particular groups of students do not have optimal opportunities to learn mathematics. Not acceptable is, for instance, that teaching methods would have more to offer to one group of students than to another one. This is not only
unacceptable from the point of view of the society, but can also not be neglected from a didactical point of view. Although it is clear that the gender differences in mathematics achievements in Dutch primary schools are not that large as the differences between children with Dutch-born parents and not Dutch-born parents — which makes that the latter research has a higher social priority — this is no reason to stop the gender study. For the further development of the RME theory, and for goals and the teaching methods that belong to it, the gender difference that was found cannot be ignored. Continued research is needed for finding the source of this difference, or, more precisely formulated, further research must make clear what education can contribute to this difference.

6 How equally suited is RME for boys and girls?

More specifically, the question can be asked what the results found in stage I of the MOOJ Project and in the PPON studies mean for RME, the approach to mathematics education that is pursued in The Netherlands. The first impression the results gave, is that the girls do better in straightforward, mechanistic mathematics education in which a lot of attention is paid to algorithms. This is just the opposite of what is aimed at in RME. The boys, on the contrary, have higher scores on problems which are close to what is in view within RME. The consequence of this could be that RME has more to offer to boys than to girls. At the same time, however, the boys could be better in lane for this approach to mathematics education: they could have more referential knowledge from daily-life, a greater arsenal of informal strategies, and could be more daring. Although not talking about RME, this opinion about a possible advantage for the boys in mathematics education was recently also stated by Geary (1997). According to him the differences between the mathematics achievements between girls and boys are larger in the advantage of boys as the quality of the education is poorer. In other words, the better the education — what is supposed to be the case in RME — the better the achievements of the boys.

For the time being, however, the contrasting results from Dutch studies into the effect of textbook series on gender differences in mathematics achievements do not confirm this hypothesis. For instance, the study done by Steen (1996) is not in agreement with additional analysis done by the PPON researchers. Steen found that the differences in mathematics scores between girls and boys were less for the students who had been taught with the realistic textbook series Pluspunt than for the students who did not work with this textbook series. But in the additional analysis of the PPON data no differential effect was found of the use of a particular textbook series on the mathematics performance of girls and boys (Bokhove, Eggen, and Van der Schoot, in preparation). Only for one subscale, written algorithms for multiplication, a significant interaction was found between textbook series and gender. For this subscale, the textbook series Operatief Rekenen, which neither can be classified as a mechanistic textbook series nor as a realistic one, turned out to be better for girls than for boys.

Other evidence that the previously stated hypothesis cannot be confirmed came from another additional analysis of the PPON data. In this analysis was investigated whether there was an interaction effect between gender and the year in which the PPON study was done. Again, no interaction effect was found (Bokhove, Eggen, and Van der Schoot, in preparation). In other words, the gender differences in mathematics achievements stayed very stable from 1987 to 1992. This means that this outcome does not
give support to the idea that RME is in the advantage of the boys. There was no increase in the gender difference in mathematics achievements, whereas at the same time it can be assumed that the implementation of RME has been improved remarkably in this period.

But is this really true? Was there really a continual growth of the implementation of RME during the years 1987 through 1992. If this was not the case, an interesting alternative explanation for the Dutch gender differences in mathematics achievement are emerging. Badly implemented RME could imply that the students, to a certain degree, are left on their own and have to use their own power to make progression. Because of their “natural” RME-related abilities, the boys could be more successful in this than the girls. The girls, on the contrary, are committed to education, are more dependent of education. The application of smart strategies, the acquisition of estimation strategies, the development of daily-life knowledge of measures, and so on, should not only be pursued as important goals, but education should also give students enough learning opportunities for achieving these goals. That is why the learning opportunities which education gives to girls will be a considerable point of interest in the observations that will be carried out the next stage of the MOOJ Project.

References


primary school. The Dutch chapter of TIMSS population 1. Enschede: OCTO, Twente University.