

Improving Reading Achievements of Struggling Readers

Wim van de Grift
Inspectorate of Education of the Netherlands

Thoni Houtveen
Department of Education
Utrecht University of Applied Sciences, The Netherlands

Contents

| | |
|--|----|
| 1. Introduction and Research Question..... | 3 |
| 2. Theoretical Framework..... | 4 |
| 2.1 Homogeneous versus Heterogeneous Ability Groups..... | 4 |
| 2.2 Offering Sufficient Opportunity to Reach the Minimum Objectives..... | 6 |
| 2.3 Using Special Teaching Techniques to Improve the Achievements of Struggling Learners..... | 8 |
| 2.4 Regular Testing of Progress and Taking Special Measures for Students who Tend to Fall Behind..... | 9 |
| 3. Research Design..... | 10 |
| 3.1 Tests Used..... | 10 |
| 3.2 Pre, Post and Follow Up Measurements..... | 11 |
| 3.3 Research Instruments for Measuring the Implementation of the Treatment..... | 11 |
| 3.4 Training of Observers and Test Assistants..... | 12 |
| 4. Findings..... | 12 |
| 4.1 Differences Between the Experimental Group and the Control Group at the Start of the Experiment..... | 13 |
| 4.2 Implementation of the Treatment..... | 13 |
| 4.4 Differences Between the Experimental Group and the Control Group one Year after the Experiment..... | 15 |
| 5. Conclusions..... | 15 |
| Literature..... | 16 |

Requests should be sent to: dr. W. J. C. M. van de Grift, Coordinating Inspector, Inspectorate of Education of the Netherlands, P.O. Box 2730, 3500 GS Utrecht, The Netherlands. Email: W.vandeGrift@owinsp.nl. Phone +31 6 22 25 29 51

Improving Reading Achievements of Struggling Readers

Abstract

The central research question is: Can we improve the results of struggling readers with optimizing their opportunities to reach the minimum targets for reading, using teaching techniques specially designed to instruct struggling readers, regular testing of students' achievements and taking special measures for students who tend to fall behind?

An experiment was carried out with 6/7-year-old children in 37 elementary schools. 21 schools with 567 children took part in the experimental group and 16 schools with 454 6/7-year-old children took part in the control group. The teachers in the experimental group have been trained in: Offering students (and especially struggling learners) ample opportunity to reach the minimum objectives for reading, using special teaching techniques to improve the achievements of struggling learners and regular testing of progress with reliable standardised tests and taking special measures for students who tend to fall behind.

At the start of the experiment the amount of students lagging behind on decoding one-syllable words in the experimental group was double the amount of students in the control group. After the experiment the situation was opposite: The amount of students lagging behind on decoding one syllable words in the control group was double the amount of students in the experimental group. After the experiment the average student in the experimental group also reached a higher reading level in a decoding test for reading whole sentences. Even one year later, when the students had other teachers these effects sustained in a moderate degree.

It is found that optimising instruction time, providing extra instruction time for struggling readers learners, setting minimum targets for reading, intensifying instruction by high interaction activities, offering 'direct instruction', adapting the assignments to the relevant differences between students, testing and diagnosing reading problems and implementing learning plans for students at risk explained the differences between the students in the experimental group and the control group.

Improving Reading Achievements of Struggling Readers

1. Introduction and Research Question

Problems of struggling readers seem to start already in elementary education with decoding. At the end of elementary education, about 10% of the students lag behind for three years in decoding. The decoding results of another 15 percent of the students at the end of elementary education lag behind for two years (Sijtsma, Van der Schoot & Hemker, 2002). These students are not able to read classical reading books especially made for their age. Most problems with decoding start already in the early years of elementary education: 10% of the 6-year-old students is not able to read in one minute more than 18 one syllable words without mistakes, while most of their peers in the classroom can already read twice as many words of one syllable in one minute (Houtveen & Van de Grift, 2001; 2007). Of course, one year later, these students also reach this criterion (the ability to read in one minute at least 18 one syllable words without mistakes). Their peers in the classroom, however have than already started with reading comprehension and a new arrear has risen for these same students.

Our hypothesis is that there should be done much more in adapting teaching to the needs of struggling readers. In a international comparative study of the inspectorates of education in England, Belgium (Flanders), Germany (Lower Saxony) and the Netherlands was found that the average scores in adapting teaching to the different needs of students of teachers in Belgium, Germany and The Netherlands did not differ significantly, but the average scores of the teachers in these three countries on adaptive teaching were lower than the average scores of the teachers in England (Van de Grift, 2007). So improvement seems possible for the Dutch teachers. The Inspectorate of Education of the Netherlands (2007) observed that teachers in schools with no struggling readers at all, scored 10 to 15 percent higher on 'differentiating instruction to the needs of students' and 'taking special measures for struggling readers than teachers in schools with struggling readers'. This is confirmed in a study of Houtveen and Van de Grift (2007). They found in a survey study on almost 1,400 students of 63 elementary schools that learning gains in decoding in classrooms with 6-year old students differ substantially. In half a year, in some classes students reached learning gains of 35 words; in other classes during half a year learning gains were no more than 10 words. In the same study, Houtveen and Van de Grift found that 'direct instruction', 'efficient organization of the instruction' and 'regular testing of achievements and taking measures for students who tend to fall behind' was significant related with high learning gains in reading (decoding).

The findings of these studies inspired to set up a quasi-experiment in which the teachers of the experimental group were specially trained in the kind of teaching needed for struggling readers. The central research question of this quasi-experiment is: Can we improve the results of struggling readers with optimizing the opportunities to reach the minimum targets for reading, using teaching techniques

specially designed to instruct struggling readers, regular testing of students' achievements and taking special measures for students who tend to fall behind?

2. Theoretical Framework

Central in the theoretical framework of this study stands the idea of 'convergent differentiation': Students lagging behind should be educated with their peers and their teachers should work with 'whole class teaching' using special techniques to make sure that even struggling learners can reach minimum standards without losing too much time.

2.1 Homogeneous versus Heterogeneous Ability Groups

There has been a lot of debate about the effectiveness of mixed versus homogeneous ability groups. Wilkinson, Hattie, Parr, Townsend, Fung, Ussher, Thrupp, Lauder and Robinson (2000) found in their meta-study in elementary and secondary education an average effect size of .10 in favour of homogeneous ability groups. However, the effectiveness of mixed and homogeneous ability groups seems to differ in elementary and secondary education. Spade, Vanfossen & Jones (1985) found, in a study of 4,000 students in secondary schools, a low and positive relationship (.10) between streaming of students and mathematics achievement. Brimer, Madaus, Chapman, Kellaghan & Woodroff (1976) found, in a study involving 44 secondary schools, low positive and high negative relationships between mixed ability groups and exam results. Kulik and Kulik (1982) found in a meta-study of 52 original studies in secondary education an average effect size of .10. In a meta-study of 29 original studies in secondary education, Slavin (1990a,b) found an average effect size of -.02.

The findings seem to be more consistent in elementary education. Kulik and Kulik (1982) found in a meta-study of 31 original studies in elementary education an average effect size of .19 in favour of homogeneous ability groups. More precise analyses made clear that this effect has to be explained by homogeneous high ability groups. For low ability groups the opposite was true. In elementary schools, heterogeneous groups appear to provide the best opportunity to learn for both low-achieving students and average students (Gamoran, 1992; Hallam & Toutounji, 1996; Houtveen & Van de Grift, 2001; Oakes, Gamoran & Page, 1992; Reezigt, 1993; Slavin, 1987; 1996). In elementary education high quality instruction given to the whole class seems to be essential for average and struggling learners.

Probably because of the confusing differential effects in elementary and secondary education, or because of the differential effects of homogeneous grouping for struggling and successful learners, struggling readers are in lots of elementary schools still segregated from their more successful classmates. Several strategies varying from 'streaming and tracking' to 'individualised instruction' and 'remedial teaching outside the classroom' are still used in trying to decrease the arrears of struggling learners. The problem is not that these techniques are wrong in itself; the problem is that these techniques often go along with side effects that resemble homogeneous grouping.

Streaming and Tracking

In some schools streaming and tracking has 'iron' rules and procedures. Poorly performing students already fall within the lowest level group in their early years and stay there throughout the rest of their time in elementary education. As a consequence, they scarcely experience any encouragement from students who are performing better. The important thing is that these schools work with a grade- or ability-level system in which not enough students can proceed to a higher level. This immobility between level groups seems to give students almost the permission to stay behind instead of helping them to increase their ability.

Individualised Instruction

At first glance, individualised instruction seems to have solved an important problem by employing an individual learning strategy, as the poorly performing student is following his or her own customized learning program. This form of adaptive teaching, however, is not appropriate for amelioration of student learning problems. Individualised instruction may foster continued poor performance, as the poorly performing student in such a program does not receive the kind of encouragement that results from being part of a larger, heterogeneous group of students. Instead, struggling readers remain behind, working in isolation at their own levels. In reality, the demands made on these students are consistently lowered. The problem is not so much the fact that individualised instruction is used, but rather that in individualised instruction insufficient attention is being paid to whether poorly performing students are getting the encouragement they need to work at a similar level to that of their peers.

Remedial Teaching Outside the Classroom

During reading lessons, struggling learners are often taken out of class for remedial teaching without the remedial support being explicitly intended to get these students performing again to the level of their peers and reintegrated into the mainstream classroom. Struggling learners often get in this kind of remedial teaching another didactic approach than the one used for all other students in the classroom. This may seem like a quite logical approach. However, the problem with this is that especially struggling learners can become quite confused because different problem-solving methods are used, and they have to work with another solution during remedial teaching than during their regular classes.

Divergent ways of differentiation like 'streaming and tracking', 'individualised instruction' and 'remedial teaching outside the classroom' may work to keep struggling learners away from closing the achievement gap with their higher performing peers. In a study of Van de Grift and Houtveen (2007) was found that about 57% of the underperforming schools in the Netherlands have too many students lagging behind schedule coupled with streaming and tracking, individualised instruction or remedial teaching outside the classroom.

Convergent Differentiation

Several educational experts like Allington and Walmsley (1995, 29) explicitly promote that 'children should be educated with their peers'. Braddock and Slavin (1995, 7) give several reasons 'why ability grouping must end'. It seems more important to work in the direction of inclusion, or convergent differentiation. In elementary education high quality instruction given to the whole class seems to be essential for average and struggling learners. However 'whole class instruction' alone is certainly not enough. Especially in classrooms with students lagging behind special teaching techniques should be used. Several reviews of the basic features of effective teaching make clear that offering students sufficient opportunity to reach the minimum objectives for reading instruction and using teaching techniques such as 'direct instruction' and offering struggling learners extra time and extra instruction lead to better learning gains of struggling learners. This should be completed with regularly testing of the progress of students and taking special measures when students tend to fall behind. This theoretical framework will be clarified in more detail in the next paragraphs.

2.2 Offering Sufficient Opportunity to Reach the Minimum Objectives

Offering students sufficient opportunity to reach the minimum objectives for reading instruction might be reached by setting minimum targets for reading instruction, planning enough time to reach these targets and making sure that planned time is efficiently used by task engaged students.

Setting Minimum Targets for Reading Instruction

Especially instruction for struggling learners asks for a target-oriented approach that works toward achieving minimum targets that have been described in detail for all students, by adapting teaching content, teaching methods, instruction time, and learning time on a needs basis indicated by test results (Carr & Harris, 2001; Corno & Snow, 1986; Slavin, Madden, Karweit, Dolan & Wasik, 1992; Houtveen, Van de Grift & Creemers, 2004; Wang, 1992).

Planning Enough Time for Reading

Planned instruction time is, seen from an international perspective, relatively high for Dutch 7 to 8 year old students. The annual load for them is 940 hours. In Belgium (Flanders) the annual load is 835, in Germany 627 and in England the annual load is 880 hours for the 7 to 8 year old students (OECD, 2007). Not only the annual load is high for Dutch students, but also the amount of time devoted to 'reading, writing and literature'. On average 30 percent of the annual load of the Dutch students in the first years of elementary education is devoted to 'reading, writing and literature'. For Belgium (Flanders), Germany and England these percentages are respectively: 22, 21 and 27 (OECD, 2007). These figures are fairly high for Dutch students. In spite of this high average percentage, schools seem to differ very much on the amount of time devoted to reading. Some schools offer their 7-8 year old students no more than 150 weekly minutes for reading activities, other schools offer their students even more than 700 minutes weekly.

Efficient Organisation of Instruction

With respect to time not only the amount of planned instruction time seems to be essential, but also the efficient use of time. Several ways of increasing allocated time are identified: beginning and ending lessons on time, reducing transition time, minimising wasted time, ensuring there are no queues waiting at the teacher's desk, well-structured lessons and orderly lesson progress (Bloom, 1976; Carroll, 1963; Harnischfeger & Wiley, 1978; and Wyne & Stuck, 1982). Efficient classroom organisation ensures that teachers do not lose much time during lesson changes and on classroom management and maintaining order (Borg, 1980; Creemers, 1994; Harnischfeger & Wiley, 1978; Kindsvatter, Willen & Ishler, 1988; Rosenshine, 1980; Rosenshine & Berliner, 1978; Wang, Reynolds & Walberg, 1993). In particular, lessons need to be efficiently organised and well-structured for poorly performing students. The importance of efficient organisation has been demonstrated by numerous researchers, such as Anderson, Evertson, and Brophy (1979), Carnine, Dixon & Silbert (1998), Dixon, Carnine, Lee & Walin (1998), Flanders (1970), Houtveen et al. (1999), and Scheerens & Bosker (1997). Several researchers developed instruments for observing efficient use of learning and teaching time. Evertson & Anderson (1978) used time logs and time sampling instruments were used by Flanders (1961), Veenman, Lem, Voeten, Winkelmoen & Lassche (1986), Roelofs (1993), The Netherlands Inspectorate of Education (1998) and Houtveen, Booiij, De Jong & Van de Grift, 1999) developed event sampling instruments. Significant correlations between time spent and students' results were found in a large number of empirical studies. (See: Scheerens & Bosker, 1997 for an overview).

Intensifying Instruction

Planned instruction time, even efficiently used, won't do if students are not engaged. Several researchers found that the degree of student engagement during an allocated period of time is of major concern (Rosenshine & Berliner, 1978; Denham & Lieberman, 1980; Fisher, Berliner, Filby, Marliave, Cahen & Denshaw (1980). Denham & Lieberman (1980) found empirical evidence that the degree to which students are engaged during allocated periods of time was directly and positively related to their learning outcomes. There is evidence that teachers can influence students' engaged learning time by avoiding excessive seatwork and by making lessons more interactive (Rosenshine, 1980; Rosenshine & Berliner, 1978). Intensifying instruction also contributes to learning gains. Anderson, Evertson & Brophy (1979) found that the percentage of academic interactions in which the student gave the correct answer was positively related to learning gains. In a later study Evertson, Anderson, Anderson & Brophy (1980) found that the use of class discussion by teachers and teacher acceptance of student ideas and contributions were more related to achievement in mathematics than to achievement in mother tongue learning. Rosenshine (1980) found that a student's task engagement rate appears to increase when the teacher directs activities actively. Task engagement during seatwork is optimised when the teacher interacts substantively with students (Rosenshine, 1980). Instruction can be intensified with the right balance in activities. When individual work is excessive, student engagement may decrease

(Rosenshine & Berliner, 1978; Rosenshine, 1980). The effectiveness of teaching increases when the lesson is characterized by a clear structure and there is good alternation between explanation, presentation, assisted tasks, independent work, and work done in small groups and with individual students (Borg, 1980; Creemers, 1994; Kindsvatter Kindsvatter, Willen & Ishler, 1988; Mortimore, Sammons, Stoll, Lewis & Ecob, 1988; Rosenshine, 1980; Rosenshine & Berliner, 1978).

2.3 Using Special Teaching Techniques to Improve the Achievements of Struggling Learners

However offering sufficient opportunity to reach the minimum objectives of the curriculum is not enough. Especially in classrooms with students lagging behind special teaching techniques should be used.

'Direct' instruction

Various authors have formulated guidelines for how to teach poorly performing students (Anderson, 1989; Rosenshine & Stevens, 1986). This guidelines are called 'direct' or 'explicit' teaching. Briefly, direct instruction involves:

- Establishing a good connection to what students already know or can do;
- Setting clear learning targets, so students know what they should know and be able to do by the end of the lesson, which is checked as well;
- Instruction content that has been broken down into small steps and is supported by specific examples and assisted tasks;
- Interactive lessons with a lot of questions and answers, as well as checks to ensure that students understand the learning material and perform the tasks correctly, together with direct feedback when this is not the case.

Essentially, these guidelines state that a lot of attention must be paid to the start of lessons, as well as to how information is transferred to students. It is important that lesson objectives are clarified at the beginning of the lesson. Students should know what they are expected to learn in the next hour. Clear specification of lesson objectives to students (Melton, 1978) and the use of clear and simple language (Land, 1987) contribute to instructional quality. Mortimore, Sammons, Stoll, Lewis & Ecob (1988) found in a study of 50 elementary schools that a limited number of focal points in each lesson were positively related to learning achievement.

Students' comprehension can be improved at the beginning of a lesson if teachers activate the required prior knowledge or, if there is none, provide students with the required information before teaching commences. Teachers can also mention core concepts, which can be used as a stepping-stone to structure information provided during the lesson. New information is transferred after the preparation stage of the lesson, which requires that the lesson aims are clear, that the lesson has a step-by-step structure, that the teacher makes connections (e.g., cause, effect, means, end of a certain event or idea), and that specific examples are used. In particular, it is important during long presentations that teachers occasionally point out core points, as well as when transitions occur, so as to simplify the structure of the learning material for students. Finally, it is very important that the teacher checks whether the new learning material has been understood (Kameenui & Carnine,

1998; Lohman, 1985; Pearson & Fielding, 1991; Pearson & Gallagher, 1983; Rosenshine & Meister, 1997).

Extra Time and Extra Help for Struggling Learners

Heterogeneous grouping is, however, certainly not enough to help those students who are at risk of school failure. These students require extended learning and instruction time. In all cases, the extension of instruction time for struggling learners demands a classroom organisation in which the remainder of the students are able to manage their own learning process. This can be done with pre-teaching and re-teaching procedures. In former years research on teaching for struggling learners was characterised by techniques for individualised instruction, streaming and remedial teaching. It has become more and more clear that individualised teaching and streaming might be prosperous for bright students, but it is disadvantageous for struggling learners, because when struggling learners work too long on their own low level they become less and less stimulated to full fill higher achievements. More recent studies make clear that struggling learners get more stimulation and are more challenged in heterogeneous groups. Working with heterogeneous groups it is important that struggling learners get extra instruction and extra time to reach the lessons targets. This development is often called the transition from divergent to convergent differentiation. Several studies showed that that extra instruction and extra learning time for struggling learners by means of pre- and re-teaching has positive results for struggling learners (Aarnoutse & Weterings, 1991, 1995; Borg, 1980; Duffy, Roehler, Meloth & Vavrus, 1986; Durkin, 1978, 1979; Houtveen, Booij, De Jong & Van de Grift, 1999; Inspectie van het Onderwijs, 1993; Kindsvatter, Wilen & Ishler, 1988; Lundberg & Linnakylä, 1992; Pearson and Fielding, 1991; Rosenshine & Berliner, 1978; Van de Grift, 1994; Van Zoelen & Houtveen, 2000; Veenman et al., 1986; Weterings & Aarnoutse, 1986; Sijstra, 1997).

2.4 Regular Testing of Progress and Taking Special Measures for Students who Tend to Fall Behind

The findings of several studies make clear that the regular testing of students progress with standardised tests go along with better achievements of students (Brandsma & Knuver, 1989; Driesen & Slegers, 2000; Houtveen, 1992; Inspectie van het Onderwijs, 2003; Janssens, 1986; Mortimore, Sammons, Stoll, Lewis & Ecob, 1988; Overmaat, 1992; Pijl, Wolfgram, Piek & Foster, 1988; Thuen & Bru, 2000; Van de Grift, 1994; Van de Grift, Houtveen & Vermeulen, 1997; Van der Werf & Guldmond, 1992; Wolf, 1977; Willms & Somers, 2001). Houtveen, Van de Grift & Creemers (2004) found, in a quasi-experiment in elementary education, that achievement in mathematics increased significantly when teachers diagnosed students' academic problems through testing and implemented prescribed plans for students identified at risk. In another study of Houtveen & Van de Grift (2007) a significant positive relationship was found between students' achievements on initial reading and frequently diagnosing students' academic problems through testing and implementing prescribed plans for students identified at risk. Achievements of students seem to be better when teacher offer more specific help

to struggling learners (Brimer, Madaus, Chapman, Kellaghan & Woodroff, 1976; Gamoran, 1992; Hallam & Toutounji, 1996; Houtveen & Van de Grift, 2001; Muijs & Reynolds, 2000; Pressley, Goodchild, Fleet, Zachowski & Evans, 1989; Oakes, Gamoran & Page, 1992; Reezigt, 1993; Spade, Vanfossen & Jones, 1985; Slavin, 1987; 1996).

3. Research Design

An experiment was carried out with 6/7-year-old children, i.e. year group 3 in Dutch elementary schools. 21 schools with 567 children took part in the experimental group and 16 schools with 454 6/7-year-old children took part in the control group. The teachers in the experimental group have been trained in: Offering students (and especially struggling learners) ample opportunity to reach the minimum objectives for reading (by setting minimum targets for reading instruction, planning enough time to reach these targets and making sure that planned time is efficiently used by task engaged students).

Using special teaching techniques to improve the achievements of struggling learners (e.g. direct instruction and offering extra time and extra help for struggling learners within heterogeneous ability groups).

Regular testing of progress with reliable standardised tests and taking special measures for students who tend to fall behind.

3.1 Tests Used

Reading Tests

To measure reading (decoding) ability, two tests are used, the Three Minute Test (Verhoeven, 1995) and the AVI-test (Visser, Van Laarhoven & Ter Beek, 1998). We used the first reading sheet of the Three Minute Test. This sheet has a list of words of the type vowel-consonant ('uil'), consonant-vowel ('koe'), consonant-vowel-consonant ('pen'). During a minute the students are asked to read this list of words aloud. The number of faultless read words is the score on the test. At the end of year 3, the students are then about 7 years old, about 90% of the students is able to read at least 18 faultless words within a minute. At the end of year 4, about 90% of the students is able to read at least 48 faultless words within a minute.

The AVI-test (Visser, Van Laarhoven & Ter Beek, 1998) is a decoding test for reading whole sentences. The test consists of 9 levels. About 50% of the students reach at the end of year 6, they are then about 10 years old, AVI-level 9. AVI-level 9 means that the student is able to read in 1 minute and 40 seconds, 22 sentences with on average 10.5 words of 1.5 syllable on average, making no more than 4 mistakes. An example of such a sentence is: 'Als je die bal ziet zweven, denk je: daar moeten we zuinig op zijn.' At the end of year 3, the students are then about 7 years old, about 75% of the students reaches at least AVI-level 2. AVI-level 2 means that the student is able to read in 1 minute and 40 seconds 24 sentences with on average 5¹/₂ words of 1 syllable making no more than 4 mistakes. An example of such a sentence is: 'Jaap gooit de bal in de lucht.' At the end of year 4, about 75% of the students reaches at least AVI-level. AVI-level 4 means that the student is able to read in 1 minute and 45 seconds 21 sentences with on average 7 words of

1.2 syllable on average making no more than 5 mistakes. An example of such a sentence is: 'Hij duwt op een andere buis die al in de grond zit.'

Intelligence Test

The 'analogies' subtest of the non-verbal SON-R (Laros & Tellegen, 1991) was used to measure intelligence. This analogies test is designed to measure the abstract reasoning ability of students aged from 5½ to 17 years. The subtest consists of 21 items. The SON-R is known as 'culture free', which is important because it was also used for students of non-Dutch origin. The test meets important reliability and validity standards (Laros & Tellegen, 1991). The subtest used is standardised with an average score 100 and a standard deviation 15.

Gender, Socio-economic and Ethnic Background and Age

The teachers were asked to complete a simple questionnaire giving the gender, the socio-economic and ethnic backgrounds and the age of each student.

3.2 Pre, Post and Follow Up Measurements

The first measurement (the pre measurement) took place in December, shortly after most initial reading content had been offered to students in Year 3. Students are supposed to know all the letters and should normally be able to voice simple words by that time. The students were tested on the first sheet of the Three Minute Test and on the Analogies subtest of the SON-R.

The second measurement (the post measurement) took place in May. The students were again tested on the first sheet of the Three Minute Test and on the AVI-test. The third measurement (the follow-up measurement) took place in May, one year later. The students were again tested on the first sheet of the Three Minute Test and on the AVI-test.

3.3 Research Instruments for Measuring the Implementation of the Treatment

The first main topic in the treatment is 'Offering students (and especially struggling learners) ample opportunity to reach the minimum objectives for reading'. This will be measured with the help of 4 different instruments:

- Setting minimum targets for reading
- Planned reading time
- Efficient organisation of instruction
- Intensifying the instruction by high interaction activities

The instruments 'Setting minimum targets for reading' and 'Planned reading time' are Likert-type questionnaires. The score range on 'Setting minimum targets for reading' is standardised from 0 to 100. This standardisation is reached by dividing the sum score on the items by the maximum score. The score on 'Planned reading time' is the average amount of weekly minutes devoted to reading. The instruments 'Efficient organisation of instruction' and 'Intensifying the instruction by high interaction activities' are both observation instruments of the event sampling type which are also standardised with a score range from 0 to 100.

The second main topic in the treatment of the experimental group is 'Using special teaching techniques to improve the achievements of struggling learners'. This will be measured with 6 different instruments:

- Direct instruction
- Adapting instruction and assignments to the relevant differences between students
- Extra instruction time for struggling learners
- Extra individual help for struggling learners
- Pre teaching struggling learners
- Re teaching struggling learners

For 'Direct instruction', 'Adapting instruction and assignments to relevant differences between students' and 'Extra individual help for struggling learners', observation instruments (event sampling type) are used with a standardised range from 0 to 100. 'Extra instruction time for struggling learners' is asked with the help of a questionnaire and measured in minutes a week. 'Pre teaching struggling learners and 'Re teaching struggling learners' is measured with a Likert type questionnaire with a standardised score between 0 and 100.

Regular testing of progress with reliable standardised tests and taking special measures for students who tend to fall behind, the third main topic in the treatment will be measured with 2 Likert type questionnaires:

- Testing and diagnosing learning problems of students at risk
- Implementing prescribed learning plans for students at risk

The scores on both Likert type questionnaires are also standardised with a score range from 0 -100.

All Likert type questionnaires were pre tested in other studies and have reliabilities (Cronbach's α) of at least .70.

3.4 Training of Observers and Test Assistants

Each observer employed on the project was given two days of training on how to score using the observation instruments. Video recordings of lessons in reading and a detailed manual were used for the training. At the end of the training a video recording was scored to determine the inter-rater reliability. This was sufficient in most cases (Hubert's $\kappa > .70$). If it was not then the observer whose scores were outside the acceptable range was given extra training, until the required criterion was reached.

A detailed manual was developed for administering the tests to the students, which described what the test assistants must do and say before and during the tests. Trainers went through these manuals with the test assistants and discussed them before the measurements were taken.

4. Findings

The following findings are discussed in turn: the pre measurements, the results of the implementation of the treatment, the results of the quasi-experiment after the implementation of the treatment and the results of the follow-up study, one year

later. The findings are concluded with the results of a multilevel regression analysis.

4.1 Differences Between the Experimental Group and the Control Group at the Start of the Experiment

Table 1 shows no significant differences between the experimental group and the control group on age, intelligence and amount of students from social deprived families. There are, however, at the start of the experiment differences in the percentage of students for which language of instruction is not their native language and in reading ability between the students in the experimental group and the control group. The amount of students for which language of instruction is not their native language is in the experimental group more than twice as high than in the control group.

At the start of the experiment, the average student in the control group is able to read almost 20 words in a minute time. In the experimental group the average student is able to read a bit more than 15 words faultless in a minute time. That is an arrears with an effect size of .48

At the start of the experiment about 13 percent of the students in het control group is not able to read more than 8 three-letter-words faultless in a minute time. The amount of students in the experimental group not able to reach the criterion of 8 words in a minute is with 28 percent more than double the amount in the control group.

4.2 Implementation of the Treatment

Table 2 shows the average scores on the implementation of the treatment by the teachers in the experimental schools in comparison with the (untrained) teachers the control schools.

[Insert Table 2 about here]

Most implementation scores are measured in standardised scores ranging from 0 to 100. The teachers in the experimental group differ a lot from the teachers in the control group on:

- Testing and diagnosing learning problems of students at risk,
- Extra instruction time for struggling learners, and
- Pre teaching struggling learners.

The teachers in the control group score about 74 percent of the items positive on 'Diagnosing learning problems of students at risk', while the teachers in the experimental group score about 92 percent of these items positive. Struggling learners in the control group get each week less than half an hour extra instruction time, the struggling learners in the experimental group more than five quarters of an hour. The teachers in the control group score less than half of the items in the scale 'Adapting instruction to relevant differences between students' positive, while teachers in the experimental group score almost 60 percent of these items positive. Small differences between the teachers in the control group and the experimental group are found for:

- Target setting for reading,
- Planned instruction time,
- Direct instruction,
- Intensifying the instruction by high interaction activities,
- Adapting instruction and assignments to the relevant differences between students, and
- Extra help for struggling learners.

The teachers in the control group scored are less often (54 percent) setting targets for reading than the teachers in the experimental group (75 percent). The teachers in the experimental group planned one hour and 10 minutes more time for reading, than the teachers in the control group did. On 'Direct instruction', 'Providing interactive instruction and activities', 'Adapting instruction and assignments to the relevant differences between students', and 'Extra help for struggling learners' the teachers in the experimental group scored about 10 percent of the items more positive than the teachers in the control group.

Only small differences are found for:

- Implementing prescribed learning plans for students at risk.
- And no differences between the teachers in the experimental group and the control group are found for:
 - Organisation of teaching, and
 - Re teaching.

Differences in scores in table 2 are measured in effect sizes. Cohen (1998) calls effect sizes of .80 and more large and effect sizes of .50 and .20 respectively and small. When we concentrate on effect sizes of .20 and more we may conclude that the actual treatment consist of :

- Offering sufficient opportunity to learn the minimum objectives of the curriculum
 - Setting minimum targets for reading
 - Planned reading time
 - Intensifying the instruction by high interaction activities
- Using teaching techniques to improve the achievements of struggling learners
 - Direct instruction
 - Adapting instruction and assignments to the relevant differences between students
 - Extra instruction time for struggling learners
 - Extra individual help for struggling learners
 - Pre teaching struggling learners
- Regular testing and taking measures for students who tend to fall behind
 - Testing and diagnosing learning problems of students at risk
 - Implementing prescribed learning plans for students at risk

We can conclude that the scores of the teachers in the experimental group on these measures are sufficiently different from the teachers in the control group.

4.3 Differences Between the Experimental Group and the Control Group after the Experiment

Table 3 shows that the students in the experimental group did catch up their arrears.

[Insert Table 3 about here]

They have read in one minute of time on average score almost 46 words faultless. The students in the control group read in the same time almost 44 words. If we correct these scores for differences in pre measurement scores the experimental group outreaches the control group with more than 5 words, which is a small effect size of .28.

At the moment of pre measurement in the experimental group the amount of students lagging behind was double the amount of students in the control group. After the experiment the situation was opposite: In the control group almost 12 percent of the students was not able to read 20 faultless words in one minute. In the experimental group less than 6 percent of the students did not reach this criterion. After the experiment the average student in the control group reached at least AVI-level 2. The students in the experimental group scored almost one AVI-level better. After correction for pre measurement, age, intelligence, social economical status and ethnicity the students in the experimental group outreached the students in the control group even more than one AVI-level. The difference has an effect size of .62.

4.4 Differences Between the Experimental Group and the Control Group one Year after the Experiment

Table 4 shows that the students in the experimental group kept their head start. The differences on the raw scores on the word test and the whole sentence test were not significant, but after correction for pre measurement, age, intelligence, social economical status and ethnicity there is still a small head start of the experimental group.

[Insert Table 4 about here]

One year after the experiment, the students in the experimental group still outreach the students in the control group with almost half an AVI-level, after correction for pre measurement, age, intelligence, social economical status and ethnicity.

5. Conclusions

At the start of the experiment the amount of students lagging behind on decoding one-syllable words in the experimental group was double the amount of students in the control group.

We made sure that the teachers in the experimental group in optimized in a better way than the teachers in the control group: the opportunities for their students to reach the minimum targets for reading by setting minimum targets for reading instruction, planning enough time to reach these targets and intensifying the

interaction during instruction. We also made sure that the teachers in the experimental group made better use of teaching techniques specially designed to instruct struggling readers such as direct instruction, extra instruction time and extra help for struggling learners like pre teaching and we made sure that the teachers in the experimental group were regularly testing the students' achievements with reliable standardised tests and that they took special measures for students who tend to fall behind.

After the experiment the situation was opposite: The amount of students lagging behind on decoding one syllable words in the control group was double the amount of students in the experimental group. After the experiment the average student in the experimental group also reached a higher reading level in a decoding test for reading whole sentences.

Even one year later, when the students had other teachers these effects sustained in a moderate degree.

We like to conclude that we can improve the results of struggling readers with optimizing the opportunities to reach the minimum targets for reading, using teaching techniques specially designed to instruct struggling readers, regular testing of students' achievements and taking special measures for students who tend to fall behind. To be sure: 'Optimising' is the keyword. We would certainly not state that the teachers in the control group were not working on the opportunities to reach the minimum targets for reading for their struggling learners or did not use teaching techniques specially designed to instruct struggling readers, or did not regularly test the students' achievements and or did not take special measures for students who tend to fall behind. The important thing is that 'a bit more' of this teaching behaviour with effect sizes between a quarter till a full standard deviation makes the difference for the struggling learners.

Literature

Aarnoutse, C.A.J. & Weterings, A.C.E.M. (1995). *Onderwijs en begrijpend lezen. [Education in comprehensive reading]* Nijmegen: Vakgroep Onderwijskunde.

Anderson, C.S. (1982). The search for climate. In: *Review of Educational Research*, vol. 52, nr. 3.

Allington, R.L. & Walmsley, S.A. (editors)(1995). *No Quick Fix. Rethinking Literacy Programs in America's Elementary Schools*. New York: Teachers College Press.

Anderson, L.W., D.W. Ryan & B.J. Shapiro. *The IEA classroom environment study*. New York: Pergamon press.

Andersen, L.M., C.M. Evertson & J.E. Brophy (1979). An experimental study of effective teaching in first grade reading groups. *The elementary school journal*, 79 (193-222).

Anderson, L. M. (1989). Classroom instruction. In M. C. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 110-116). New York: Pergamon Press.

Bloom, B.S. (1976). *Human Characteristics and School Learning*. New York: Mc Graw Hill.

- Borg, W. R. (1980). Time and school learning. In C. Denham & A. Lieberman (Eds.). *Time to learn* (pp. 33-72).
- Braddock, J. H. & Slavin, R.E. (1995) Why ability grouping must end: achieving excellence and equity in American education. In H. Pool & J.A. Page, *Beyond tracking, finding success in inclusive schools* (pp. 7-19). Bloomington, Indiana: Phi Delta Kappa Education Foundation.
- Brandsma, H.P. & Knuver, J.W.M. (1989). *Basisschoolkenmerken als determinanten van het functioneren van leerlingen*. [Characteristics of elementary schools as determinants for the functioning of pupils] Groningen: RION.
- Brimer, A., G.F. Madaus, B. Chapman, T. Kellaghan & R. Woodroff (1976). *Differences in School Achievement*. Slauch: NFER-Nelson.
- Carroll, J.B. (1963). A model for school learning. In: Anderson, L.W. (Ed.), *Perspectives on school learning selected writings of John B. Carroll*, pp. 19-31.
- Carnine, D.W. (1979). Direct instruction: A successful system for educationally high-risk children. *Journal of Curriculum Studies*, 11 (1), 29-45.
- Carnine, D.W., Dixon, R.C. & Silbert, J. (1998). Effective Strategies for Teaching Mathematics. In: Kameenui, E.J. & Carnine, D.W. (eds). *Effective Teaching Strategies That Accommodate Diverse Learners*. New Jersey: Prentice Hall.
- Carr, J. F. & Harris, D. E. (2001). *Succeeding with standards. Linking curriculum, assessment, and action planning*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*, 1st Edition, Lawrence Erlbaum Associates, Hillsdale (2nd Edition, 1988).
- Corno, L., & Snow, R.E. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Ed.). *Handbook of research on teaching* (3rd ed., pp. 605-629). New York: Macmillan Publishing Co.
- Creemers, B.P.M. (1994) *The Effective Classroom*. London: Cassell Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Denham, C. & Lieberman, A. (1980) *Time to learn*. Washington, DC: National Institute of Education.
- Dixon, R., Carnine, D.W., Lee, D.W. & Wallin, J. (1998). *Review of high quality experimental mathematics research*. Austin: University of Texas.
- Duffy, G.G., Roehler, L.R. Sivan, E., Rackliffe, G. , Book, C. , Meloth, M.S. , Vavrus, L.G. Wesselman, R., Putnam, J., Bassiri, D. (1986). Effects of Explaining the Reasoning Associated with Using Reading Strategies *Reading Research Quarterly*, Vol. 22, No. 3 (Summer, 1987), pp. 347-368
- Durkin, D. (1978-1979). What classroom observation reveals about reading comprehension instruction. *Reading Research Quarterly*, 14, 481-533.
- Driesen & Slegers, 2000
- Evertson, C.M. & Anderson, L.M. (1978). *Interim Progress Report: the Classroom Organization Study*. Austin: University of Texas
- Evertson, C.M., C.W. Anderson, L.M. Anderson & J.E. Brophy (1980). Relationships between classroom behaviors and student outcomes in junior high mathematics and English classes. *American Educational Research Journal*, 17 (1), 43-60.

- Fisher, C. W., Berliner, D. G., Filby, N. N., Marliane, R., Cahen, L. S. & Denshaw, M. (1980). Teaching behaviors, academic learning times, and student achievement: An overview. In C. Denharn & A. Lieberman (Eds.). *Time to learn* (pp. 7-32). Washington DC: National Institute of Education.
- Flanders, N. (1970). *Analyzing teaching behavior*. Reading (Mass.): Addison-Wesley.
- Gamoran, A. (1992). Is ability grouping equitable: Synthesis of research. *Educational Leadership*, 50 (1), 11-17.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Charles Griffin & Company Ltd.
- Hallam, S. & Toutounji, I. (1996). *What do we know about the ability grouping of pupils by ability? A research review*. London: Institute of Education, University of London.
- Haquebord, H.I., Linthorst, T.R., Stellingwerf, B.P. & De Zeeuw, M. (2004). *Voortgezet taalvaardig* [Continuity in language competences] Groningen: ETOC.
- Harnishfeger, A. & Wiley, D.E. (1978). Conceptual issues in models of school learning. *Curriculum Studies*, 10, 215-131.
- Heij, K., Haitjema, T. & Lam, J.F. (2008). *Taalniveau in het mbo* [Language competencies in vocational training]. Lienden: ICE.
- Houtveen, A.A.M., N. Booij, R. de Jong & W.J.C.M. van de Grift (1999). Adaptive Instruction and Pupil Achievement. *School Effectiveness and School Improvement*, 10 (2) 172-192.
- Houtveen, A.A.M. & W.J.C.M. van de Grift (2001). Inclusion and Adaptive Instruction in Elementary Education *Journal of Education for Students Placed At Risk*, 6 (4) 389-411.
- Houtveen, A.A.M., W.J.C.M. van de Grift & B.P.M. Creemers (2004). Effective School Improvement in Mathematics. *School Effectiveness and School Improvement*, 15 (3-4) 337-376.
- Houtveen, A.A.M. & W.J.C.M. van de Grift (2007). Effects of Meta cognitive Strategy Instruction and Instruction Time on Reading Comprehension. *School Effectiveness and School Improvement*, 18 (2) 173-190.
- Houtveen, A.A.M., W.J.C.M. van de Grift, J. Kuijpers, M. Boot, F. Groot & H. Kooijman (2007). Improving Underperforming Schools Working on Improvement. *Journal of Education for Students Placed At Risk*, 12 (4) 361-381.
- Houtveen, A.A.M. & W.J.C.M. van de Grift (2007). Reading Instruction for Struggling learners *Journal of Education for Students Placed At Risk*, 12 (4) 405-424.
- Mullis, I.V.S., Martin, M.O., Kennedy, A.M. & Foy, P. (2007). *Progress in international Reading Literacy Study in Primary School in 40 Countries*. Chestnut Hill, MA.: TIMSS & PIRLS International Study Center, Boston College.
- Inspectie van het Onderwijs (2007). *The state of education in the Netherlands 2005/2006*. Utrecht: Inspectie van het Onderwijs
- Janssens, F.J.G. (1986). *De evaluatiepraktijken van leerkrachten: een beschrijvend onderzoek naar het evalueren tijdens het rekenen in het primair onderwijs*. [Monitoring of pupils' results in primary education] Arnhem: CITO.

- Kameenui, E.J. & Carnine, D.W. (Eds.).(1998). *Effective teaching strategies that accommodate diverse learners*. Columbus, OH: Merrill-Prentice Hall.
- Kindsvatter, R., Wilen, W. & Ishler, M. (1988). *Dynamics of effective teaching*. NY: Longman.
- Kulik, J.A. & Kulik, C.L. (1988). Timing of feedback and verbal learning. *Review of Educational research*, 58, 79-97.
- Land, M.L. (1987). Vagueness and clarity. In: Dunkin, M.J. (Ed.). *International Encyclopedia of Teaching and Teacher Education*. New York: Pergamon.
- Laros, J.A. & Telligen, P.J. (1991). Construction and validation of the SONR 5-17, The Snijders-Oomen non-verbal intelligence test. Groningen: Wolters-Noordhoff.
- Lohman, D. F. (1985). *Teaching higher-order skills*. Elmhurst, NJ: North Central Laboratory for Educational Research and Development.
- Lundberg & Linnakylä, 1992. *Teaching reading around the world*. The Hague: IEA.
- Melton, R. F. (1978). Resolution of conflicting claims concerning the effects of behavioural objectives on student learning. *Review of Educational Research*, 48, 291-302.
- Mortimore, P., P. Sammons, L. Stoll, D. Lewis & R. Ecob (1988). *School matters: the junior years*. Somerset: Open books
- Muijs, D. & Reynolds, D. (2000). School effectiveness and teacher effectiveness: Some preliminary findings from the evaluation of the mathematics enhancement programme. *School Effectiveness and School Improvement*, 11, 247-263.
- Oakes, J., Gamoran, A. & Page, R.N. (1992). Curriculum differentiation: Opportunities, outcomes and meanings. In P.W. Jackson (Ed.), *Handbook of Research on Curriculum*. (pp. 570-609). Washington DC: AERA.
- OECD (2007). *PISA 2006, Science competencies for tomorrows world. Vol. 1 & 2*. Paris: OECD.
- OECD (2007). *Education at a glance*. Paris: OECD.
- Overmaat, A.M. (1992). *De kwaliteit van meetinstrumenten voor het aanbod van taal en lezen*. [Quality of Measurement Instruments for Reading and Writing.] Amsterdam: SCO.
- Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research* (Volume 2, pp. 815-860). New York: Longman.
- Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, 8, 317-344.
- Pijl, S.J., Wolfgram, H.P., Piek, J. & S.F. Foster (1988). *Voortgangsevaluatie bij het lees- en spellingsonderwijs* [Evaluation of Reading and Spelling.] Groningen: RION.
- Pressley, M., Goodchild, J. Fleet, R., Zachowski, R. & Evans, E. (1989). The challenges of classroom strategy instruction. *Elementary School Journal*, 58, 266-278.
- Reezigt, G.J. (1993). *Effecten van differentiatie op de basisschool* [Effects of differentiation in primary education]. Groningen: RION.
- Roelofs, E. C. (1993). *Teamgerichte nascholing en coaching. Een experimentele studie in scholen met combinatieklassen*. [Staff development

- and coaching: An experimental study in schools with combined classes.]. Nijmegen, The Netherlands: Katholieke Universiteit.
- Rosenshine, B. (1976). Classroom instruction. In: N. Gage (eds.) *The psychology of teaching methods*. National society for the study of teaching (335-371).
- Rosenshine, B.V. (1986). Synthesis of research on explicit teaching. *Educational Leadership*, 60-69.
- Rosenshine, B.V. & Stevens, R. (1986). Teaching functions. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). New York: Macmillan.
- Rosenshine, B.V. & Meister, C. (1997). Cognitive Strategy Instruction in Reading. In: Stahl, S. & Hayes, D.A. (eds) *Instructional Models in Reading*. Mahwah, New Jersey: The Guilfords Press.
- Rosenshine, B. (1979). Content, time and direct instruction. In: Peterson, P.L. & H.J. Walberg (Eds.), *Research on teaching*. Berkeley: McCutchan.
- Rosenshine, B.V. & B.C. Berliner (1978). Academic engaged time. *British Journal of Teacher Education*, 4, 3-16.
- Rosenshine, B. & J. Furst (1971). Current and future research in teacher performance and criteria. In: Smith, B.W. (Ed.) *Research on teacher education*. Englewood Cliffs: Prentice Hall.
- Rosenshine, B. & R. Stevens (1986). Teaching functions. In: Wittrock, M.C. (Ed.). *Handbook of research on teaching*, third edition. New York: MacMillan.
- Rosenshine, B. (1979). Content, time and direct instruction. In: Peterson, P.L. & H.J. Walberg (Eds.), *Research on teaching*. Berkeley: McCutchan.
- Rosenshine, B.V. & B.C. Berliner (1978). Academic engaged time. *British Journal of Teacher Education*, 4, 3-16.
- Rosenshine, B. & J. Furst (1971). Current and future research in teacher performance and criteria. In: Smith, B.W. (Ed.) *Research on teacher education*. Englewood Cliffs: Prentice Hall.
- Rosenshine, B. V. & Berliner, D. C. (1978). Academic engaged time. *British Journal of Teacher Education*, 3, 316.
- Rosenshine, B. (1980). How time is spent in elementary classrooms. In C. Denham & A. Lieberman (Eds). *Time to learn* (pp. 107-126). Washington, DC: National Institute of Education.
- Rosenshine, B. (1983). Teaching functions in instructional programs. *Elementary School Journal*, 83(4), 335-351.
- Rosenshine, B. & R. Stevens (1986). Teaching functions. In: Wittrock, M.C. (Ed.). *Handbook of research on teaching*, third edition. New York: MacMillan.
- Scheerens, J. & Bosker, R. (1997). *The Foundations of Educational Effectiveness*. Oxford: Pergamon Press.
- Sijtsma, J., F. van der Schoot & B. Hemker (2002). *Balans van het taalonderwijs aan het einde van de basisschool 3. Uitkomsten van de derde peiling in 1998*. [Reading achievements at the end of primary education] Arnhem: Citogroep. Insitituut voor Toetsontwikkeling.
- Slavin, R. (1990). Achievement effects of ability grouping in secondary schools: A best evidence synthesis. *Review of Educational Research*, 60, (3), 471-499.
- Slavin, R. (1990). Ability grouping in secondary schools: A response to Hallinan. *Review of Educational Research*, 60, (3), 505-507.

- Slavin, R.E. (1989). A theory of school and classroom organization. In: Slavin, R.E. (1989) *School and classroom organization*. New Jersey: Lawrence Erlbaum Associates Inc.
- Slavin, R.E. (1987). Ability grouping and achievement in elementary schools. *Review of educational research*, 57, 293-336.
- Slavin, R.E. (1996). *Education for all. Contexts of learning*. Lisse: Swets & Zeitlinger.
- Slavin, R., Madden, N. A., Karweit, N. L., Dolan, L., & Wasik, B.A. (1992). *Success for all: A relentless approach to prevention and early intervention in elementary schools*. Arlington, VA : Educational Research Service.
- Spade, J.Z., B.E. Vanfossen & E.D. Jones (1985). *Effective schools: Characteristics of schools which predict mathematics and science performance*. Chicago: AERA.
- Van de Grift, W., Th. Houtveen & C. Vermeulen (1997). Instructional Climate in Dutch Secondary Schools. *School Effectiveness and School Improvement*, 8 (4) 449-462.
- Van de Grift, W. & A.A.M. Houtveen, (2006). Underperformance in Primary Schools. *School Effectiveness and School Improvement*, 17 (3) 255-273.
- Van de Grift, W. (2007). Quality of Teaching in Four European Countries: a review of the literature and an application of an assessment instrument. *Educational Research*, 49 (2) 127-152.
- Van de Grift, W. & A.A.M. Houtveen, (2007). Weaknesses in Underperforming Schools. *Journal of Education for Students Placed At Risk*, 12 (4) 383-403.
- Veenman, Lem, Voeten, Winkelmolen & Lassche (1986),
- Wang, M. C., Reynolds, M. C., & Walberg, H. J. (1993). Serving students at the margin. *Educational Leadership*, 52(4), 15-29.
- Wilkinson, I., Hattie, J., Parr, J., Townsend, M., Fung, I., Ussher, C., Thrupp, M., Lauder, H., & Robinson, T. (2000). *Influence of peer effects on learning outcomes: A review of the literature*. Auckland, New Zealand: Auckland UniServices Limited.
- Veenman, S., Lem, P., Voeten, B., Winkelmolen, B. & H. Lassche (1986). *Onderwijs in combinatieklassen*. [Education in multi graded class rooms] 's-Gravenhage: SVO.
- Verhoeven, L. (1995). *Drie-minuten-toets* [Three Minute Test]. Arnhem: Cito.
- Visser, J., Van Laarhoven, A., & Ter Beek, A. (1994). *AVI-toetspakket*. [AVI-test] 's-Hertogenbosch: Katholiek Pedagogisch Centrum.
- Willms, J.D. & Somers, M.A. (2001). Family, Classroom and School effects on Children's Educational Outcomes in Latin America. *School Effectiveness and School Improvement*, 12 (4), 409-445.
- Wijnstra, J.M. (2001). *Bruikbare kennis en vaardigheden voor jonge mensen*. [Useful knowledge and skills for young people] Arnhem: Cito
- Wolf, R.G. (1977). *Achievement in America*. New York: Teachers College Press.
- Van Zoelen, E.M. & Houtveen, A.A.M. (2000). *Naar effectieve schoolverbetering*. [Towards Effective School Improvement]. Utrecht: ISOR/Onderwijsonderzoek.

Wyne, M. D. & Stuck, G. B. (1982). Time and learning: Implications for the classroom teacher. *Elementary School Journal*, 83(1), 67-75.

Table 1. Pre measurements in December

| | Control group N=454 students N=16 schools | Experimental group N=567 students N=21 schools | Effect size (Cohen's d) |
|---|---|--|----------------------------|
| Age (in months) | 87.97 | 87.77 | -.02 |
| Intelligence (SON-R-test) | 99.87 | 100.06 | .01 |
| % students from social deprived families | 12.67 | 12.65 | -.06 |
| % students for which language of instruction is not their native language | 6.82 | 13.87 | .21 |
| Average score decoding single words (Sheet 1 DMT-test) | 19.61 | 15.51 | -.30 |
| % students unable to read more than 8 words in one minute | 13.35 | 28.30 | .48 |

Table 2. Measurements about the implementation of the treatment

| | Control group N=16 schools | Experimental group N=21 schools | Effect size (Cohen's d) |
|---|----------------------------------|---------------------------------------|----------------------------|
| Offering sufficient opportunity to learn the minimum objectives of the curriculum | | | |
| Setting minimum targets for reading (range: 0 -100) | 54.16 | 74.70 | .73 |
| Planned reading time (minutes per week; range: 150 – 750) | 362.20 | 432.15 | .54 |
| Efficient organisation of instruction (range: 0 -100) | 84.05 | 82.97 | -.07 |
| Intensifying the instruction by high interaction activities (range: 0 -100) | 48.19 | 56.66 | .42 |
| Using teaching techniques to improve the achievements of struggling learners | | | |
| Direct instruction (range: 0 -100) | 50.28 | 59.06 | .52 |
| Adapting instruction and assignments to the relevant differences between students (range: 0 –100) | 43.39 | 54.04 | .59 |
| Extra instruction time for struggling learners (minutes per week; range: 0 - 180) | 27.55 | 76.84 | .99 |
| Extra individual help for struggling learners (range: 0 -100) | 65.67 | 75.03 | .50 |
| Pre teaching struggling learners (range: 0 -100) | 49.95 | 59.32 | .88 |
| Re teaching struggling learners (range: 0 -100) | 75.03 | 76.76 | .09 |
| Regular testing and taking measures for students who tend to fall behind | | | |
| Testing and diagnosing learning problems of students at risk (range: 0-100) | 74.47 | 91.18 | .95 |
| Implementing prescribed learning plans for students at risk (range: 0 - 100) | 88.25 | 93.45 | .27 |

Table 3. Post measurements in May, a half year later

| | Control group N=454 students N=16 schools | Experimental group N=567 students N=21 schools | Effect size (Cohen's d) |
|--|---|--|----------------------------|
| Post measurements in May | | | |
| Average score decoding single words (Sheet 1 DMT-test) | 43.59 | 45.83 | .11 |
| % students unable to read more than 20 words in one minute | 11.61 | 5.74 | -.18 |
| Average score reading whole sentences (AVI-test) | 2.25 | 3.11 | .46 |
| Post measurements in May corrected for pre measurement, age, intelligence, social economic status and ethnicity | | | |
| Average score decoding (Sheet 1 DMT-test) | 41.87 | 47.42 | .28 |
| Average score reading whole sentences (AVI-test) | 2.09 | 3.26 | .62 |

Table 4. Follow up measurements in May, one and a half year later

| | Control group N=454 students N=16 schools | Experimental group N=567 students N=21 schools | Effect size (Cohen's d) |
|---|---|--|----------------------------|
| Follow up measurements in May, one and a half year later | | | |
| Average score decoding (Sheet 1 DMT-test) | 69.10 | 70.89 | .09 |
| Average score reading whole sentences (AV-I-test) | 5.86 | 6.18 | .16 |
| Follow up measurements in May corrected for pre measurement, age, intelligence, social economic status and ethnicity | | | |
| Average score decoding (Sheet 1 DMT-test) | 68.88 | 72.42 | .18 |
| Average score reading whole sentences (AVI-test) | 5.83 | 6.32 | .24 |