School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/nses20

Effects of metacognitive strategy instruction and instruction time on reading comprehension

A. A. M. Houtveen [a] & W. J. C. M. van de Grift [b]

[a] Utrecht University of Applied Sciences , The Netherlands

[b] Dutch Inspectorate of Education , Soestduinen, The Netherlands

Published online: 29 May 2007.

To cite this article: A. A. M. Houtveen & W. J. C. M. van de Grift (2007) Effects of metacognitive strategy instruction and instruction time on reading comprehension, School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice, 18:2, 173-190, DOI: 10.1080/09243450601058717

To link to this article: http://dx.doi.org/10.1080/09243450601058717

School Effectiveness and School Improvement
Vol. 18, No. 2, June 2007, pp. 173–190

Effects of Metacognitive Strategy Instruction and Instruction Time on Reading Comprehension

A. A. M. Houtveen[a]* and W. J. C. M. van de Grift[b]
[a]*Utrecht University of Applied Sciences, The Netherlands,* [b]*Dutch Inspectorate of Education, Soestduinen, The Netherlands*

(Received 30 May 2006; accepted 7 September 2006)

Ten-year-old students in an experimental group and a control group were tested for metacognitive abilities in reading comprehension before and after implementation of treatment in the experimental group. The teachers of the students in the experimental group were trained in metacognitive strategy instruction and in optimizing instruction time for reading comprehension. The learning gains made by the students in the experimental group in metacognitive abilities for reading comprehension turned out to be significantly greater than those made by the students in the control group. In the next school, year the students in the former experimental group and the former control group were tested for reading comprehension. It turned out that the students in the former experimental group had significantly better results on reading comprehension than the students in the former control group.

Introduction and the Research Question

About 10 years ago, international comparative research revealed that Dutch 9-year-olds had average scores for reading comprehension of narrative, persuasive, and informative texts of 494, 480, and 481, respectively. The equivalent scores achieved by Finnish students, who led the international league table, were 568, 569, and 569. That is a substantial difference, with a standard deviation of about 85 points (Elley, 1992). This relatively poor performance by the Dutch students inspired us to start a quasi-experiment to find answers to the question: ''What can teachers do to improve their students' performance in reading comprehension?''

*Corresponding author. Department of Education, Utrecht University of Applied Sciences, PO Box 14007, 3508 SB Utrecht, The Netherlands. Email: thoni.houtveen@hu.nl

ISSN 0924-3453 (print)/ISSN 1744-5124 (online)/07/020173–18

DOI: 10.1080/09243450601058717

Explanations for poor performance in reading comprehension have been sought in a number of areas, including the influence of reading outside school, characteristics of schools, the reading materials used in school, qualities of groups and teachers, and the teaching methods employed (Lundberg & Linnakylä, 1992; Postlethwaite & Ross, 1992). Ten years ago, many educators believed that reading comprehension had to be "caught rather than taught" (Pearson, Roehler, Dole, & Duffy, 1992), and they seemed to have reasons to back up this way of thinking. Before 1986, there were no reviews of research into reading comprehension instruction (Pearson & Fielding, 1991). Durkin (1979) was probably the first who made perfectly clear that there was virtually no instruction in reading comprehension, with teachers spending less than 1% of their time on instruction during reading comprehension lessons. The limited time spent on reading comprehension and the virtually complete absence of metacognitive strategy instruction in Dutch schools have also been pointed out (Aarnoutse & Weterings, 1995; Weterings & Aarnoutse, 1986).

While it may be true that the influence of teachers on the student performance is less than the influence of students' intelligence and socio economic background, teachers do still have a real influence. After all, various studies into elementary education have shown that between 6 and 26% of the differences in student achievement levels can be influenced by factors in the school and classroom (Bosker & Witziers, 1996; Brandsma & Knuver, 1989; Roeleveld, 2003; Wijnstra, Ouwens, & Béguin, 2003). These percentages mark the operational scope of the potential effects of teaching. This article reports on the findings of a quasi-experiment into the effects of metacognitive strategy instruction and optimizing instruction time.

A Model of Comprehension Instruction

Much work on the process of reading comprehension has been grounded in studies of good readers (Block & Pressley, 2001; Duke & Pearson, 2002; Pearson et al., 1992; Pressley & Afflerbach, 1995). In summary, these authors stated that good readers:

- are active readers;
- constantly evaluate whether the text, and their reading of it, is meeting their goals;
- look over the text before they read, noting the structure of the text and text sections that might be most relevant to their reading goals;
- make frequent predictions about what is to come;
- read selectively and continually make decisions about their reading (what to read carefully, what to read quickly, what not to read, what to reread, and so on);
- construct, revise, and question the meanings they make as they read;
- try to determine the meaning of unfamiliar words and concepts in the text and deal with inconsistencies or gaps as needed;
- draw from, compare, and integrate their prior knowledge with material in the text;
- monitor their understanding of the text.

Since the 1990s, researchers have suggested that the strategies identified in the behaviour of good readers should form the basis for comprehension instruction. Pearson and Fielding (1991), Graves, Watts, and Graves (1994), Rosenshine and Meister (1997), Gersten, Fuchs, Williams, and Baker (2001), Collins Block and Pressley (2002), and Duke and Pearson (2002) have distinguished the following aspects of instruction based on reviews of good and expert readers.

Before Reading

Before a student starts to read a text, it is important that he learns to perform a number of activities. These are mainly concerned with deciding on the purpose of reading the text and activating prior knowledge. It is also important that the teacher prepares the students for their reading of the text in an active and focused way, by predicting the content based on characteristics of the text and its structure. A rough scan of the text is especially necessary to gain a general impression of the theme or content of the text, to establish which parts of the text are important in view of the purpose of reading, to assess what is known and what is not known, and to decide what strategies are needed. In order to gain an impression of the content of a text quickly, knowledge of the specific structure of both narrative and informative texts is required. This means knowing about how texts are structured into chapters, sections and paragraphs, components of texts such as the title, subheadings, summary, use of punctuation, and layout. After the reader has formed an impression of the content of the text, he can predict what its content will be and what he can expect.

During Reading

During the reading of a text, the reader needs to learn to use the following strategies and skills: checking understanding of what he is reading; restoring understanding of what he is reading; verification and possible adjustment of the first impression of the text; distinguishing between main topics and secondary topics; gathering information from the text; deducing the meaning of unfamiliar words from the context; being able to use an index. The following teaching principles are important when teaching these strategies and skills: explicit explanation of the strategy/skill; demonstrating use of the strategy/skill whilst explaining the process aloud (acting as a model); offering support (''scaffolds'') in the use of the strategy/skill (reciprocal teaching appears to be especially valuable for reading comprehension); motivating students to apply the strategy/skill; guiding the students as they learn to master the strategy or skill through practice, so that responsibility is increasingly shifted on to the students during the course of the learning process. Demonstration (modelling) is seen as one of the most useful techniques for explaining strategies and skills for reading comprehension. By talking through the strategy or skill, the teacher (or a good student) shows how it should be done. Making it explicit in this way helps poor readers by making clear what they should be doing and what they were not doing before, or what they were doing wrong (Aarnoutse, 1991; Rosenshine & Meister, 1997).

After Reading

After reading a text (and doing any assignments on it), it is important that the teacher checks what the students have understood and gives them feedback on their use of the strategies and skills. Conclusions must be drawn both on the text itself and on the students' use of the strategies and skills. A teacher will do well to emphasize where the appropriate use of a strategy led to the correct answer (attribution of success to use of the strategy). Recently, the importance of interaction to concept formation has also been emphasized. Pearson (2001) argued that reading comprehension should be placed in a context in which those involved discuss ideas, feelings, and insights linked to a clear learning objective. Honig (2001) emphasized the importance of involving students in in-depth discussion of texts. Students must be given the opportunity and skills to discuss the text and the use of strategies with fellow students and the teacher.

By the end of the 1990s, there was a great deal of evidence that students can be taught some of the strategies used by good readers and that when they are taught those strategies their understanding and memory of texts read improves. Many experiments were conducted in which one group of students was taught to use some particular strategy while reading (e.g., summarization), with their reading compared with that of students who were not given this instruction. In general, the students taught to use the comprehension strategy outperformed the students not given instruction. A number of individual strategies were validated as effective in improving comprehension in students in the upper elementary grades. There is fairly clear evidence that upper-elementary students can be taught to relate what they are reading to prior knowledge, to construct mental images of text content, to question themselves about ideas in the text, and to summarize what has been read. There is much less evidence that we know how to teach students to make inferences and to monitor their text comprehension (Pressley, 2002). In most studies, the strategies were taught one at a time (Keene & Zimmermann, 1997). Since it is recognized that good readers articulate multiple strategies as they read, it seems sensible to teach young readers to coordinate multiple strategies as they read. At the moment, there is quite a bit of evidence that elementary students benefit from instruction in the use of a small repertoire of reading comprehension strategies (Pressley, 2002).

In the research literature, teaching of comprehension strategies has followed a common model recently, beginning with teacher modelling and explanation, and explanation of strategies, and continuing with student practice of strategies with the teacher, then coaching students' use of strategies (Anderson, 1989; Duke & Pearson, 2002; Kameenui & Carnine, 1998; Rosenshine, 1986; Rosenshine & Stevens, 1986; Simmons & Kameenui, 1998). Not only does such teaching increase reading comprehension, it does so by encouraging the kind of active reading done by proficient readers. This model of comprehension instruction can be broken down into the following five stages:

1. an explicit description of the strategy and when and how it should be used;
2. teacher and/or student modelling of the strategy in action;

3. collaborative use of the strategy in action;
4. guided practice using the strategy with gradual release of responsibility;
5. independent use of the strategy. (Anderson, 1989; Duke & Pearson, 2002; Kameenui & Carnine, 1998; Rosenshine, 1986; Rosenshine & Stevens, 1986; Simmons & Kameenui, 1998).

Throughout these five stages, it is important that neither the teachers nor the students lose sight of the need to coordinate or orchestrate comprehension strategies. Strategies are not to be used singly. Although the above description of the model emphasizes a particular strategy at a particular time, other strategies should also be referenced, modelled, and encouraged throughout the process.

A great deal of attention must be paid to the initial phases of the lesson and the way information is transferred. The students' understanding can be boosted at the start of the lesson if the teachers activate the necessary prior knowledge or, in the absence of any prior knowledge, teach this before the students start to read. After this initial preparatory phase of the lesson, new information can be transferred. It is important in this phase that the objectives of the lesson are clear and that there is a phased structure to the lesson, that the teacher makes connections (cause, effect, means, objective of an event or idea), and that concrete examples are given. A strategy is explained in explicit terms. The teacher models it, that is, he or she demonstrates the use of the strategy. This explicit instruction method makes clear to the students what the strategy entails and when it should be used. This is metacognitive information about the strategy. The strategy serves, in other words, as a stepping-stone for the structuring of information that will be dealt with in the lesson. In the phase of guided practice to learn to master the strategy, the teacher gives careful feedback to the students on how to improve implementation of the strategy. The teacher praises and rewards the students and corrects them where necessary. Having them apply the strategy in different contexts encourages transfer. The students then apply the strategy themselves during the independent assimilation phase, and in this phase greater use is made of forms of instruction geared to self-regulation, encouraging the students to reflect on their own cognitive skills and the promotion of autonomous learning (Pearson & Fielding, 1991; Rosenshine & Meister, 1997).

A way of conceptualizing the orchestration process is captured in the classic visual model of Pearson and Gallagher (1983). In that model, teachers move from a situation in which they assume all the responsibility for performing a task while the student assumes none to a situation in which students assume all the responsibility while the teacher assumes none. At the start of the lesson, the teacher demonstrates or models a strategy. Instruction at that time would be labelled teacher centred. At the end of the lesson, the student practices the strategy independently. Instruction then would be labelled student centred. The phase of guided practice to learn to master the strategy lies between these two extremes. In this phase, the teacher's responsibility gradually decreases and the student's responsibility gradually increases. In this phase, the teacher provides scaffolds to the students. The concept of ''providing scaffolds'' is closely related to the concept of the ''zone of proximal development'' (Vygotsky,

1978). The use of scaffolding enables students to solve problems that they would not have been able to solve without this support. The student gets support from the teacher to complete tasks that would be just too difficult for him or her without help (tasks which lie in the student's zone of proximal development). By using scaffolding, the teacher reduces the complexity of the situation, gives structure, clarifies the problem, points the student in the direction of the next step he or she has to take, monitors the purpose of the activity, involves the student in completion of a joint task, creates a framework, and gives rules that the student gradually adopts himself (self-regulation). Examples of scaffolds are reciprocal teaching, giving procedural support (such as an action guide), acting as a model, demonstrating and at the same time explaining the procedure aloud (or getting a good student to do this), and having the students work in mixed ability groups. Acting as a model and demonstrating, whilst at the same time explaining the procedure aloud, are considered to be particularly important. These forms of support are withdrawn as soon as the student can complete the task on his or her own.

Instruction Time for Reading Comprehension

Instruction time and time management are important at both the school and class levels. The classroom time planned for language and reading varies in Dutch classrooms around 7 hr per week (Sijtstra, 1997; Van de Grift, 1994). Teachers do not generally spend much time on teaching reading comprehension. According to Pearson and Fielding (1991), teachers let their students read texts for between 7 and 15 min a day, by which they mean times when the main issue is not correct decoding of the text but understanding the content of the text. Not much time is set aside for reading comprehension therefore.

The planned 7 hr for language and reading already mentioned is not, of course, net "language and reading time." Time is lost by moving between lessons, by class management, and by maintaining discipline. Teachers differ in the way they manage time in the classroom. Van de Grift (1994) found that on average about 44% of lesson time was spent on teaching, 50% on independent work, and 6% on class management. About 85% of the teachers needed less than 10% of their time for class management. (Houtveen, Booij, De Jong, & Van de Grift, 1999; Van Zoelen & Houtveen, 2000; Veenman, Lem, Voeten, Winkelmolen, & Lassche, 1986). This research also showed that efficient teachers spend 15% less time on class management and organizational tasks than less efficient teachers and that they spend 50% more time on interactive activities with students.

Less able students, and younger students in particular, would benefit from the teaching time spent on reading comprehension being increased. After all, a number of studies have shown that the amount of time spent on teaching is directly related to educational achievement (Aarnoutse & Weterings, 1991; Lundberg & Linnakylä, 1992; Rosenshine & Berliner, 1978; Weterings & Aarnoutse, 1986).

Of course, it is not only the amount of learning time that is important. Getting the right balance between the components of a lesson is also important. If left to work

independently for too long, students lose their concentration (Rosenshine, 1980; Rosenshine & Berliner, 1978). Efficient teachers organize their time so that they not only have time for the whole group but also for working with small groups or individual students (Borg, 1980; Kindsvatter, Wilen, & Ishler, 1988).

There are further special requirements for time management in reading comprehension lessons. Aarnoutse and Weterings found in various studies (Aarnoutse & Weterings, 1991, 1995; Weterings & Aarnoutse, 1986) that, in lessons on reading comprehension, students were only instructed in how to actually apply skills and strategies for a fraction of the lesson time. Pearson and Fielding (1991) also found that the time the teacher spent in reading comprehension lessons on explicitly explaining the use of strategies was minimal. Students were given hardly any help in learning to resolve problems adequately while reading by having skills and strategies explained and demonstrated to them (Duffy, 1983; Duffy, Roehler, Meloth, & Vavrus, 1986; Durkin, 1979; Mason & Osborn, 1983). It appears, therefore, that the time spent on teaching skills and strategies is limited. In lessons devoted to reading comprehension, teachers are mainly occupied with organizational aspects of the lesson, such as giving out books, asking questions, and checking answers (Aarnoutse & Weterings, 1991, 1995; Durkin, 1979; Pressley et al., 1998; Weterings & Aarnoutse, 1986).

Research Design

A quasi-experiment was carried out with 10-year-old children, that is, Year 6 in Dutch elementary schools. Eleven schools with 344 children took part in the experimental group, and nine schools with 225 10-year-old children took part in the control group. About 19% of the children in both groups were from low socio-economic status (SES) native families and about 2.5% of the children in both groups were from low-SES ethnic minority families. The teachers in the experimental group were trained in cognitive strategy instruction and in optimizing teaching time for reading comprehension.

Training of Teachers

To allow the experiment to be conducted, it was necessary to find schools that were willing to implement the treatment. It was vital in this situation that the experimental variables be implemented as faithfully as possible. This called for the necessary supervision and control of the conduct of the experiment. In order to bring this about, a contract was signed between the researchers and the school advisory services, in which the activities of and time to be spent by those involved were laid down. An improvement programme for reading comprehension was written for this purpose. This programme included the content of the innovation, a supervision programme, and a general plan laying down which parts of the treatment must be discussed and when. This programme had to be implemented. After all, if the supervisors did not perform their supervisory activities properly, the teachers could not be expected to be

able to realize the treatment in their classes. The following criteria were used to judge whether implementation of the supervision strategy was adequate:

- at least 1 half-day per 14 days must be invested in the project;
- the supervisory activities must be directed at all levels in the school: school principal, teams, and individual teachers of the 10-year-old children;
- attention must be given to all subjects set out in the project plan;
- there must be a mix of supervisory interventions, with an emphasis on interventions geared to improving teaching behaviour.

Based on the situation in the schools, the supervisors were able to decide for themselves, to a certain extent, in what order to deal with the different subjects and the degree of emphasis given to each subject in a particular school. A monitoring system was used to chart implementation of the planned supervisory activities. By comparing the activities carried out with the planned activities, they were able to find out how far the plans had been realized. The monitoring system consisted of logbook record forms on which supervisors reported all the supervisory activities they had carried out as part of the project. The supervisors completed one of these forms after each supervision meeting. The completed forms were then used to find out whether the supervision strategy had been realized to an adequate extent. This was possible because the criteria for adequate supervision were included on the forms as notes for the attention of the supervisors.

The school supervisors had the option not to do the classroom visits themselves but to coach the school principal to carry out the classroom consultations. A classroom consultation meant observing the teacher in the class followed by a discussion. Where the school principal had taken on responsibility for the classroom consultations, this was recorded, as the school principal filled in the form in that case. By adding up or ranking these data for the whole supervision period, it was possible to test whether the supervisors had fulfilled each of the criteria. Of course, the fact that the supervision strategy had been carried out properly did not necessarily mean that the teachers had actually implemented the treatment. This point will now be examined.

Measuring Metacognitive Strategy Instruction

In order to establish the extent to which teachers were working along the lines of the model for teaching metacognitive strategy instruction in reading comprehension, we used an observation instrument specifically developed for this research. The measurement objective was: ''Extent of working with the strategic action model for metacognitive strategy instruction.'' This observation instrument used the ''event-sampling'' procedure. This procedure is suitable for assessing aspects in the classroom situation that are not really amenable to measurements over time. The instrument consists of 20 items on teacher behaviour. The homogeneity (Cronbach's α) of the instrument varied between .80 and .94. The inter-rater reliability (Hubert's

κ) varied between .70 and .78 (Booij, Houtveen, & Overmars, 1995; Houtveen & Booij, 1994). The following findings emerged from a generalizability study in which eight observers used this instrument on the same teachers on two measurement occasions. The unreliability caused by differences between the observers was 14% and the unreliability caused by differences in measurement occasions was virtually 0%. The unreliability caused by the interaction of observers and measurement occasions was 3% in this generalizability study. The remaining variance (83%) was explained by differences between the observed persons or interactions with persons (Booij et al., 1995).

Measuring Instruction Time

A variety of instruments is available for measuring teaching time spent on reading comprehension (Barr & Sadow, 1989; Durkin, 1979; Sindelar, Smith, Harriman, Hale, & Wilson, 1986; Wendler, Samuals, & Moore, 1989; Weterings & Aarnoutse, 1986). These related instruments inspired the development of a new observation instrument, ''Teaching time for reading comprehension.'' This is a time-sampling instrument. It scores according to the predominant activity principle, that is to say that the activity that was dominant (occurred the most) in the interval (7 s) is scored during the 13 s afterwards. The inter-rater reliability (Hubert's κ) varied between .80 and .84 (Booij et al., 1995; Houtveen & Booij, 1994).

Questionnaire for Metacognitive Knowledge

The reading comprehension questionnaire developed by Gruwel and Aarnoutse (1995) aims to measure the metacognitive knowledge that students have about reading strategies. It concerns strategies that a student can use to check his own reading behaviour (monitoring) and strategies that a student can use when he cannot understand a text or paragraph. Six types of questions are asked in the questionnaire. The questions relate to strategies that can be used before, during, and after reading, strategies for evaluating the reading process, and strategies that can be used to resolve things that are unclear. The questionnaire consists of 30 multiple choice questions, and is administered to the whole class at once. The homogeneity of the questionnaire varied between .78 and .84 on different administration occasions (Gruwel & Aarnoutse, 1995). This is more than satisfactory.

Questionnaire on Reading Attitude and Reading Material

Reading attitude was measured using a reading attitude scale published by Berkhout Publishing (Aarnoutse, 1991). This scale attempts to measure students' attitudes to reading and reading materials. Reading attitude is taken to mean the tendency or disposition of students to value reading situations and reading material positively or negatively. The scale contains a total of 27 statements. To counteract the tendency to

habit forming in responses, there are 17 positively worded questions and 10 negatively worded questions. The homogeneity of the scale is more than satisfactory at .90 (Aarnoutse, 1991).

Test for Measuring Reading Comprehension

The "Test for Measuring Reading Comprehension" from the student monitoring system of the Netherlands Institute for Educational Measurement (Cito) (Staphorsius & Krom, 1998) was used to measure the students' performance in reading comprehension in the follow-up measurement. This test has been evaluated by the Commission on Testing Matters (COTAN) of the Netherlands Institute of Psychologists (NIP) and was judged to be "good" on the quality of the principles on which the tests were constructed, the quality of the test material, the quality of the manual, reliability, and construct validity (Resing et al., 2002).

Gender, Socioeconomic and Ethnic Background, and Age

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

Intelligence Test

The "analogies" subtest of the nonverbal SON-R (Laros & Tellegen, 1991) was used to measure intelligence. The Snijders – Oomen Nonverbal Intelligence Test (SON-R) is a test for the individual assessment of intelligence in which the use of spoken or written language is not necessary. The test was originally designed for assessment of deaf children. The 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 a test with a low cultural loading, 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).

Training of Observers and Test Assistants

Each observer employed on the project was given 4 half-days of training on how to score using the observation instruments. Video recordings of lessons in reading comprehension 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.

Findings

The following findings are discussed in turn: the results of implementation of the treatment, the results of the quasi-experiment, and the results of the follow-up study.

Implementation of the Treatment

Table 1 shows the scores on implementation of the treatment by the teachers in the experimental schools, in comparison with the teachers in the control schools, who had not been trained in the treatment.

Table 1 shows that the scores of the teachers in the experimental group outperformed the control group on metacognitive strategy instruction and instruction time with effect sizes (Cohen's *d*) of .34 and .87, respectively. Using Cohen's (1988) classification, the effect sizes found can be said to be medium and large, respectively. We can conclude that the scores of the teachers in the experimental group are sufficiently different from the teachers in the control group to expect differences in student results.

Table 1. Average scores in standard scores, years, and percentages

	Control group[a]	Experimental group[b]	Effect size (Cohen's *d*)
Premeasurements			
Premeasurement metacognitive skills	10.37	11.09	.12
Intelligence	99.56	100.12	.04
Reading attitude	17.36	18.88	.25
Age (in years)	8.26	8.17	.24
Girls (%)	51.20	50.17	
Native pupils with low SES (%)	23.56	16.28	
Ethnic minority pupils with low SES (%)	4.44	1.16	
Postmeasurements			
Postmeasurement metacognitive skills (uncorrected)	15.02	17.32	.38
Postmeasurement metacognitive skills (corrected for premeasurement, intelligence, reading attitude, low SES, ethnic minority, gender, and age)	15.26	17.40	.36
Implementation measurements			
Metacognitive strategy instruction	45.40	50.06	.34
Instruction time	25.01	34.61	.87

[a] $n = 225$ pupils, $n = 9$ schools.
[b] $n = 344$ pupils, $n = 11$ schools.

Differences in Metacognitive Knowledge

Table 1 shows that there were substantial and significant postmeasurement differences between the cognitive performance of the 10-year-old students in the experimental and control groups. For metacognitive knowledge, the effect size in favour of the experimental group was .38, which is a medium difference.

Explanation of the Differences Between Experimental and Control Groups

Table 1 shows only very minor differences between the experimental group and the control group on premeasurement scores on metacognitive knowledge, intelligence, and gender. On age, reading attitude, and socioeconomic and ethnic background, however, some differences were found. Differences in reading attitude and socioeconomic background especially are expected to parallel differences in metacognitive knowledge and reading comprehension. So, we corrected for the differences in premeasurement intelligence, reading attitude, low SES, ethnic minority status, gender, and age in order to find out how much of the differences in metacognitive knowledge can be explained by the treatment. When we corrected (with ANOVA) the postmeasurement scores on metacognitive knowledge for these differences between the control group and the experimental group, the scores were 15.26 and 17.40, respectively. The effect size of the difference between the control group and the experimental group was .36, which is still a moderate difference.

Table 2. Results of the multilevel analysis ($n = 569$ pupils, 20 schools, 2 conditions)

Explained variance in metacognitive skills	Pupils	School	Condition
Explained variance in the "0 model"	.758	.075	.131
Variance to be explained after introduction of all variables	.586	.006	.000
	β	SE	
Pupil level			
Premeasurement metacognitive skills	.31	.05[a]	
Intelligence	.18	.05[a]	
Girl	.18	.10	
Age	−.11	.09	
Native pupil with low SES	−.28	.12[a]	
Low SES pupil from ethnic minority	−.04	.39	
Reading attitude	.14	.05[a]	
School/condition level			
Metacognitive strategy instruction	.30	.06[a]	
Instruction time	.33	.07[a]	

[a]Significant at .05 level.

Multilevel analysis was used to analyse whether the experimental effect could be explained by the experimental treatment. We used the MLWin programme for this procedure. The results are presented in Table 2.

We computed the amount of variance explained by the individual level and the school level and the level of the experimental or control condition on the posttest. It turned out that about 79% of the variance in student results on the posttest could be explained by differences between individual students. About 21% was school/classroom variance. About two thirds of this school/classroom variance was due to differences between the experimental group and the control group.

The next step aimed at explaining the differences between students that were not due to school or the experiment. The student background variables scores on the pretest, gender, intelligence, age, reading attitude, and socioeconomic and ethnic background were added to the model. Only the pretest score, intelligence, reading attitude, and socioeconomic background were found to be significantly related to posttest results. No significant differences were found between boys and girls and no effects were found for ethnic minority students. Taken together, these variables were able to explain about 23% of the student variance.

In the final model, we added the treatment variables ''metacognitive strategy instruction'' and ''instruction time.'' The variance between schools not explained by differences due to the experimental and control condition reduced from .075 to .006. Furthermore, it turned out that all of the variance in student results that could be explained by differences in the experimental conditions was, indeed, explained by these experimental variables. The variance due to differences between the experimental and the control condition was reduced to zero.

Findings of the Follow-Up Study

The next year, almost all the students from the control group and the experimental group transferred to other classrooms, usually with other teachers. During the first months of the next school year, all students from both the control group and the experimental group were tested on reading comprehension. The results are shown in Table 3.

Table 3. Average standard scores in reading comprehension in the follow-up study

	Control group[a]	Experimental group[b]	Effect size (Cohen's *d*)
Reading comprehension (uncorrected)	32.30	36.28	.44
Reading comprehension (corrected for premeasurement on metacognitive skills, intelligence, reading attitude, low SES, ethnic minority, gender, and age)	33.08	36.39	.36

[a] $n = 225$ pupils, $n = 9$ schools.
[b] $n = 344$ pupils, $n = 11$ schools.

The average score of the students in the control group was 32.30 and the average score of the students in the experimental group was 36.28. The difference shows a medium size effect of .44. After correction (with ANOVA) for the differences on premeasurement scores on metacognitive knowledge, intelligence, gender, reading attitude, socioeconomic and ethnic backgrounds, and age between the experimental group and the control group, the average scores were 33.08 and 36.39 for the control and experimental groups, respectively. The effect size of the difference between the control group and the experimental group after correction was .36, which is still moderate and in favour of the experimental group.

Conclusions and Discussion

We established that the teachers in the experimental group demonstrated better ''metacognitive strategy instruction'' and devoted more instruction time to their students than the teachers in the control group. We also found that the students in the experimental group made greater progress in metacognitive knowledge than their contemporaries in the control group. Can we conclude from this that better ''metacognitive strategy instruction'' and ''more instruction time'' accounts for these learning gains in metacognitive skills? The non-equivalent control group design that we used is robust enough to control for a number of alternative explanations for the experimental effect found. Working with a control group allowed us to rule out the assumption that special events or the natural development and increased maturity of the students could provide an alternative explanation for the experimental effects. The use of pretest and posttest tests for all groups in the experiments allowed us to also rule out differential effects of pretest and posttest as an alternative explanation. The measurement instruments used were not changed and the same instruments were used for both the control and experimental groups, so they cannot have played a role either. ANOVA analysis and multilevel regression analysis was used to control for statistical regression, differential selection of respondents, and the problem of interaction effects. Of course, it is possible to debate the question as to whether we used the right control variables. However, in our opinion, by choosing gender, intelligence, age, reading attitude, and socio-economic and ethnic background we had covered all the obvious control variables. Despite the fact that our choice of control variables could be criticized, we still feel able to conclude that better ''metacognitive strategy instruction'' and ''more dedicated instruction time'' did lead to an improvement in cognitive performance.

From the follow-up study we learned that the students in the former experimental group had significantly better results on reading comprehension than the students in the former control group. We believe that an improvement in metacognitive skills leads to better results in reading comprehension. Although many teachers and even some researchers still seem to think that achievements in reading comprehension are mostly based on innate abilities like intelligence, our findings make it clear that there is at least some evidence that reading comprehension can be learned and

taught. Reading comprehension is not a matter of unchangeable and innate abilities. Teachers can teach their students metacognitive knowledge such as:

- activating prior knowledge;
- using the title, subheadings, the summary, punctuation, and layout to predict the content of the text;
- making frequent predictions about what is to come;
- reading selectively and making decisions about the reading process (what to read carefully, what to read quickly, what not to read, what to reread, and so on);
- drawing from, comparing, and integrating prior knowledge with material in the text;
- monitoring their understanding of the text;
- checking their understanding of the content.

Teachers who engage in specific teaching activities will have pupils with better achievements in reading comprehension. It is important that teachers:

- explicitly explain the above-mentioned strategies and skills;
- demonstrate use of these strategies and skills and act as a model;
- offer support (scaffolds) in the use of these strategies and skills;
- motivate students to apply strategies and skills;
- guide students as they learn to master the strategy or skill;
- give students increasing self-responsibility during the learning process;
- check what the students have understood;
- give students feedback on their use of the strategies and skills;
- let students draw conclusions on the text and their use of strategies;
- emphasize, where appropriate, that strategies led to the correct answer;
- place reading comprehension in an authentic context;
- discuss ideas and insights linked with the learning objective;
- involve students in in-depth discussions of texts;
- give students opportunities to discuss the text and use of the strategy.

References

Aarnoutse, C. A. J. (1991). *Begrijpend leestest* [Test for comprehensive reading]. Nijmegen, The Netherlands: Berkhout.

Aarnoutse, C. A. J., & Weterings, A. C. E. M. (1991). *Leeractiviteiten en processen tijdens het onderwijs in begrijpend lezen* [Activities and processes during instruction in comprehensive reading]. Nijmegen, The Netherlands: Interdisciplinaire Studierichting Onderwijskunde.

Aarnoutse, C. A. J., & Weterings, A. C. E. M. (1995). Onderwijs in begrijpend lezen [Instruction in comprehensive reading]. *Pedagogische Studiën*, *72*, 82–101.

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

Barr, R., & Sadow, M. W. (1989). Influence of basal programs on fourth-grade reading instruction. *Reading Research Quarterly*, *24*(1), 44–71.

Block, C. C., & Pressley, M. (2001). *Comprehension instruction: Research-based best practices*. New York: Guilford.

Booij, N., Houtveen, A. A. M., & Overmars, A. M. (1995). *Instructie bij begrijpend lezen: Constructie van 2 observatie-instrumenten voor het meten van de kwaliteit en kwantiteit van instructie bij begrijpend lezen* [Instruction in reading comprehension]. Utrecht, The Netherlands: Universiteit Utrecht.

Bosker, R. J., & Witziers, B. (1996, April). *The magnitude of school effects, or: Does it really matter which school a student attends?* Paper presented at the Annual Meeting of the American Educational Research Association, New York.

Brandsma, H. P., & Knuver, J. W. M. (1989). Effects of school and class room characteristics on pupils progress in language ant arithmetic. *International Journal of Educational Research*, *13*, 777–788.

Borg, W. R. (1980). Time and school learning. In C. Denham & A. Lieberman (Eds.), *Time to learn* (pp. 33–72). Washington, DC: National Institute of Education.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edn). Hillsdale, NJ: Lawrence Erlbaum Associates.

Collins Block, C., & Pressley, M. (2002). *Comprehension instruction. Research-based best practices.* New York: The Guilford Press.

Duffy, G. G. (1983). *From turn taking to sense making: Classroom factors and improved reading achievement.* Lansing, MI: Michigan State University, Institute for Research on Teaching.

Duffy, G. G., Roehler, L. R., Meloth, M. S., & Vavrus, L. G. (1986). Conceptualization instructional explanation. *Teacher and Teacher Education*, *2*, 197–214.

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd edn, pp. 205–242). Newark, DE: International Reading Association.

Durkin, D. (1979). What classroom observation reveals about reading comprehension instruction. *Reading Research Quarterly*, *14*, 481–533.

Elley, W. B. (1992). *How in the world do students read?* Den Haag, The Netherlands: I.E.A.

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: Review of research. *Review of Educational Research*, *71*(2), 279–321.

Graves, M. F., Watts, S., & Graves, B. (1994). *Essentials of classroom teaching: Elementary reading* (pp. 165–191). Nietham Heights, MA: Allyn & Bacon.

Gruwel, S., & Aarnoutse, C. A. J. (1995). *Vragenlijst: Begrijpend lezen en test: Zoek de fouten* [Questionnaire on comprehensive reading]. Nijmegen, The Netherlands: Berkhout.

Honig, B. (2001). *Teaching our children to read.* Thousand Oaks, CA: Corwin Press.

Houtveen, A. A. M., & Booij, N. (1994). *Het meten van integrale leerlingzorg: Adaptief onderwijs en schoolontwikkeling* [The measurement of adaptive instruction and school development]. Utrecht, The Netherlands: Universiteit Utrecht.

Houtveen, A. A. M., Booij, N., de Jong, R., & Van de Grift, W. J. C. M. (1999). Adaptive instruction and pupil achievement. *School Effectiveness & School Improvement*, *10*(2), 172–193.

Kameenui, E. J., & Carnine, D. W. (1998). *Effective teaching strategies that accommodate diverse learners.* Columbus, OH: Merrill-Prentice Hall.

Keene, E., & Zimmerman, S. (1997). *Mosaic of thought*. Portsmouth, NH: Heinemann.

Kindsvatter, R., Wilen, W., & Ishler, M. (1988). *Dynamics of effective teaching.* New York: Longman.

Laros, J. A., & Tellegen, P. J. (1991). *Construction and validation of the SON-R 5-17, the Snijders-Oomen non-verbal intelligence test.* Groningen, The Netherlands: Wolters-Noordhoff.

Lundberg, I., & Linnakylä, P. (1992). *Teaching reading around the world.* Den Haag, The Netherlands: IEA.

Mason, J. M., & Osborn, J. (1983). *When do children begin "reading to learn": A survey of classroom reading instruction practices in grades two through five.* Urbana, IL: University of Illinois Center for the Study of Reading.

Pearson, P. D. (2001). Life in the radical middle: A personal apology for a balanced view of reading. In R. Filippo (Ed.), *Reading researchers in search of common ground* (pp. 815–860). Newark, DE: International Reading Association.

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* (Vol. 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.

Pearson, P. D., Roehler, L. R., Dole, J. A., & Duffy, G. G. (1992). Developing expertise in reading comprehension. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (2nd edn, pp. 145–199). Newark, DE: International Reading Association.

Postlethwaite, T. N., & Ross, K. N. (1992). *Effective schools in reading: implications for educational planners.* Den Haag, The Netherlands: IEA.

Pressley, M. (2002). *Reading instruction that works. The case for balanced teaching*. (2nd edn). New York/London: The Guilford Press.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.

Pressley, M., Wharton-McDonald, R., Mistretta-Hampston, J., & Echevarria, M. (1998). Literacy instruction in 10 fourth-grade classrooms in upstate New York. *Scientific Studies of Reading*, *2*(2), 159–194.

Resing, W. C. M., Evers, A., Koomen, H. M. Y., Pameijer, M. K., Bleichrodt, N., & Van Boxtel, H. (2002). *Indicatiestelling: Condities en instrumentarium in het kader van leerlinggebonden financiering* [Quality of research instruments for individual testing of children]. Lochem, The Netherlands: NDC Boom.

Roeleveld, J. (2003). *Herkomstkenmerken en begintoets. Secundaire analyses op het PRIMA-cohort onderzoek* [Social and economical backgrounds and testing]. Amsterdam: Kohnstamm Instituut.

Rosenshine, B. V. (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. V. (1986). Synthesis of research on explicit teaching. *Educational Leadership*, *43*, 60–69.

Rosenshine, B. V., & Berliner, D. (1978). Academic engaged time. *British Journal of Teacher Education*, *4*, 3–16.

Rosenshine, B. V., & Meister, C. (1997). Cognitive strategy instruction in reading. In S. Stahl & D. A. Hayes (Eds.), *Instructional models in reading.* Mahwah, NJ: The Guilford Press.

Rosenshine, B. V., & Stevens, R. (1984). Classroom instruction in reading. In P. D. Pearson (Ed.), *Recent research on reading.* New York: Longman.

Rosenshine, B. V., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd edn, pp. 376–391). New York, NY: Macmillan.

Sijtstra, J. (1997). *Balans van het taalonderwijs aan het einde van de basisschool: uitkomsten van de tweede taalpeiling einde basisonderwijs* [Achievements in the mother tongue at the end of primary education]. Arnhem, The Netherlands: Cito.

Simmons, D. C., & Kameenui, E. J. (1998). *What reading research tells us about children with diverse learning needs. Bases and basics.* Mahwah, NJ: Lawrence Erlbaum Associates.

Sindelar, P. T., Smith, M. A., Harriman, N. E., Hale, R. L., & Wilson, R. J. (1986). Teacher effectiveness in special education programs. *Journal of Special Education*, *20*(2), 195–207.

Staphorsius, G., & Krom, R. S. H. (1998). *Toetsen begrijpend lezen* [Tests on reading comprehension]. Arnhem, The Netherlands: Cito.

Van de Grift, W. J. C. M. (1994). *Technisch rapport van het onderzoek onder 386 basisscholen ten behoeve van de evaluatie van het basisonderwijs* [Report on the evaluation of 386 schools for primary education]. De Meern, The Netherlands: Inspectie van het Onderwijs.

Van Zoelen, E. M., & Houtveen, A. A. M. (2000). *Naar effectieve schoolverbetering. Resultaten van het onderzoek naar de effecten op onderwijsgevende-en leerlingniveau van het adaptief onderwijsproject "Kwaliteitsversterking Rekenen en Wiskunde"* [Effects of school improvement. An experiment on adaptive instruction in arithmetic lessons]. Utrecht, The Netherlands: ISOR/ Onderwijsonderzoek.

Veenman, S., Lem, P., Voeten, B., Winkelmolen, B., & Lassche, H. (1986). *Onderwijs in combinatieklassen* [Education in multigraded classrooms]. 's-Gravenhage, The Netherlands: SVO.

Vygotsky, L. S. (1978). *Mind in society.* Cambridge, MA: MIT Press.

Wendler, D., Samuals, S. J., & Moore, V. K. (1989). Comprehension instruction of award-winning teachers, teachers with master's degrees and other teachers. *Reading Research Quarterly*, *24*(4), 382 – 401.

Weterings, A. C. E. M., & Aarnoutse, C. A. J. (1986). De praktijk van het onderwijs in begrijpend lezen [The practice of instruction in comprehensive reading]. *Pedagogische Studiën*, *63*, 387 – 400.

Wijnstra, J., Ouwens, M., & Béguin, A. (2003). *De toegevoegde waarde van de basisschool* [Added value of schools in elementary education]. Arnhem, The Netherlands: Citogroep.