Chapter 2

Effectiveness of Guided Co-Construction versus Direct Instruction for Beginning Reading Instruction[1]

[1]Snel, M.J., Terwel, J., Aarnoutse, C.A.J. & van Leeuwe, J.F.J. (2012). Effectiveness of guided co-construction versus direct instruction for beginning reading instruction, Educational Research and Evaluation,18 (4), 353–374.

Abstract

In a field experiment with 178 first grade pupils, the effect of an experimental beginning reading programme was investigated. Both an experimental and a control group worked with the most frequently-used Dutch beginning reading programme, Learning to Read Safely. The instructional approach implemented in the experimental group was Guided Co-Construction (GCC); the instructional approach implemented in the control group was Direct Instruction (DI).

The results of an overall repeated measurement analysis of the development of word recognition (WR) over time (i.e., throughout the first grade) showed the pupils in the experimental group to outperform those in the control group. However, the better performance by the experimental group attenuated over time. The control group almost catches up to the mean of the experimental group by the fourth measurement occasion. Majority pupils benefitted more from GCC but minority pupils more from DI. Minority pupils in the control group showed greatest progress.

2.1 Introduction

Teaching children to read is a complex task. Children enter school with substantial speaking competence but little or no reading or writing skills. The purpose of beginning school reading instruction is thus to help children master the many challenges of the written word, including knowledge of the alphabetic system, an ability to decode new words and an ability to construct, integrate and remember the meanings of words in text. In order for children to link spoken language to written language, they must master the alphabetic principle or, in other words, a system of grapheme-phoneme correspondences which associate the spellings of words with their pronunciations (Ehri, 1991). There are nevertheless large differences across children in the mastery of the alphabetic principle, and the aim of this study was therefore to determine what form of beginning reading instruction facilitates children's work recognition the most: direct instruction or guided co-construction?

2.1.1 Stages in the development of reading

According to Ehri (1991) and Chall (1996), children proceed through predictable phases when beginning to read. Chall distinguishes three phases: phase 0 or prereading, which typically developing readers achieve around six years of age; phase 1 or the initial reading or decoding phase, which typically developing readers reach by six or seven years of age; and phase 2 or confirmation, which typically developing readers reach at around the age of eight. These phases are very similar to the first three phases in Ehri's model of reading development: phase 1 or the prealphabetic phase; phase 2 or the alphabetic phase, which consists of the partial and full alphabetic subphases; and phase 3 or the consolidated

alphabetic phase. The prealphabetic phase has also been called the logographic phase because it occurs before the development of alphabetic knowledge (Ehri, 1991). Children are able to recognize certain words by sight (i.e., due to distinctive visual and contextual cues around or in the recognized words). The logographic reading of cereal-box labels, restaurant logos and other types of environmental print is thus among the first literacy accomplishments of the preschool child. The reading of signs and logos shows that the young child is attending to visual cues in his or her surroundings; the young child may also attend to visual cues within words and thus read the word 'moon' by recognizing the two circles in the middle of the word.

When children develop knowledge of letters in words and specific letter-sound relationships, they enter the partial alphabetic subphase of the second phase of word recognition. This occurs during kindergarten or first grade when most children notice that particular letters in a word correspond to particular sounds in the pronunciation of the word. For example, a child may recognize 'mask' by recognizing the letter-sound relationships for the initial 'm' and the final 'k' but not for the letters in between.

Children enter the full alphabetic subphase when they can match all of the letters and sounds in the alphabet. At this phase in the development of word recognition and to actually read words, children can segment the word 'moon' into three letter units which match three pronounced sounds. Sounding out letters and blending them into words may be laborious and slow at the beginning of the full alphabetic phase but, as children become more accomplished at decoding unknown words, they progress to more rapid word analysis.

The consolidated alphabetic phase emerges when children consolidate graphemes into chunks or specific spelling patterns. With increased experience and the reinforcement of particular word patterns, children are now able to read many words and syllables on the basis of memory or via analogy to hundreds of words which share the same spelling pattern: bat, hat, cat, mat, fat, sat and so on. With practice, more words get stored in memory and recognized more or less automatically. Reading is no longer slow and analytic but, rather, fluent (Gentry, 2006). This process typically continues through fourth grade or, for poor readers, even sixth grade.

2.1.2 Socio-cultural background

According to The National Early Literacy Panel (NELP, 2008), The Dutch Inspection of Education (2006e, 2008b) and Stoep (2008), beginning reading performance — in contrast to performance in other subjects — hardly relates to the socio-economic or ethnic backgrounds of pupils. Only when pupils are asked to read particularly long or complex words a difference in performance does emerge (Droop & Verhoeven, 2003; Verhoeven, 2000): Children with lower social-economic backgrounds score lower on measures of

early decoding skills than those from upper/middle social economic backgrounds (Hecht, Burgess, Torgesen, Wagner, & Rashotte, 2000; NICHD, 2000).

Leseman and de Jong (1998) reported similar significant effects of socio-cultural background on the word decoding skills of seven-year-old children. The effects stem from three dimensions of home education (Leseman & de Jong, 1998, 2001), namely reading opportunity, instructional quality and social-emotional quality. The degree to which the home environment provides reading opportunities can obviously affect the development of word decoding skills. And while 'opportunity' refers to the quantity of a wide range of reading experiences, Leseman and de Jong (1998) further distinguished the degree and nature of parental guidance during literacy interactions with their children. And the social–emotional quality of the relationship between the parents and the child appeared to play a role as well. Minority parents are reading less with their children than majority parents do, give their children less autonomy and indicate less confidence in their interactions with their children. The strength of the relationship between home education and word decoding declined between 6 and 9 years of age, which means that the influence of home education on the development of children's word decoding is limited to the initial stages of learning to read. Minority children may thus be accustomed to initially greater step-by-step instruction like that provided by direct instruction while majority children may benefit more from a cooperative learning approach. Along these lines, Edmonds (1977), Popp and Lieberman (1977), Venezky (1978) and Weber (1971) have all shown the provision of reading and study-skill instruction to contribute significantly to the reading achievement of pupils and those whose parents have not had advanced schooling in particular.

2.1.3 The role of reading instruction

One of the major questions for theories of learning and instruction is whether knowledge should be provided or generated (see Rosenshine, Meister, & Chapman, 1996). A similar dichotomy can also be seen in the domain of reading education (Stahl, 1999). A great deal of the instructional approaches in early reading can be categorized in more 'teacher-directed' or more 'child-directed' approaches. Discussions about the best ways to teach reading in the early years are often caught up in a dilemma of stimulating spontaneous reading activities of children without explicit phonics instruction versus teacher-structured exercises of rules in which phonics are taught in order to automatically recognize words. Proponents in both camps sometimes take extreme positions, but also drive out new ideas (Stahl, 1999). Although the effectiveness of an instructional approach is always related to the learning content and learning objectives, one can generally conclude that if generating is understood as individual discovery learning, it is less effective as compared to providing methods. In addition, such a radical form of generating seems an unrealistic option in mainstream education. Indeed, learning in the classroom is a social event: teachers and fellow pupils will always have some input in the learning process of an individual pupil. 'Generation', conceived as a radical constructivist approach does not exist in normal classroom practice.

Therefore, searching for a third way seems an interesting option to overcome the dilemma. This third way is found in the approach 'guided co-construction' (GCC) of knowledge. It is a structured approach with a cooperative learning component (see also Hardman 2008; Mercer, 1995; Terwel, van Oers, van Dijk, & van den Eeden, 2009). Such an approach to instruction has also been called co-elaboration, co-construction or the guided reinvention of the language symbol system (Brown & Palincsar, 1989; Dewey, 1943; Freudenthal, 1991). In the present research we were therefore interested in the effects of two instructional approaches, a providing approach: Direct Instruction (DI) and a third way approach with a cooperative learning component: Guided Co-Construction (GCC). 'Direct instruction' (DI) or the direct provision of knowledge is known to be an effective instructional approach, particularly for children at risk for reading difficulties (Adams, 1990; Anderson, Hiebert, Scott, & Wilkinson, 1985; Bus & van IJzendoorn, 1999; Chall, 1996; Ehri, Nunes, Stahl, & Willows, 2001; Evans & Carr, 1985; Hattie, 2008; Slavin, Lake, Chambers, Cheung, & Davis, 2009; Stahl & Miller, 1989). In reviews and meta-analyses, however, Raudenbush (2009) and Slavin et al. (2009) have also shown instructional approaches which include a cooperative learning component to be effective for beginning reading instruction.

In the following, DI and GCC will be discussed in greater detail. The different phases of DI will be briefly outlined. And GCC will be shown to be particularly relevant for the teaching of beginning reading.

Direct Instruction. Direct instruction (DI) has been studied in several domains of teaching, including the instruction of mathematics and language. Within the context of beginning reading instruction, Slavin et al. (2009, p. 1406) has defined DI as: '... an approach to beginning reading instruction that emphasizes a step-by-step approach to phonics, decodable texts that make use of a unique initial teaching alphabet, and structured guides for teachers.' The instruction is highly structured and describes or even scripts classroom activities in considerable detail. The emphasis is squarely on the systematic teaching of the written language code. DI addresses both 'what' to teach (i.e., the content of a curriculum) and 'how' to teach (i.e., specific techniques).

In an analysis of those teaching behaviours and organizational factors associated with positive pupil learning outcomes, Rosenshine and Stevens (1986) identifies particularly effective instructional practices and grouped them into six phases for DI. a) Review: This phase serves to motivate pupils, to briefly summarize the previous lesson and to make the purpose of the present lesson clear. b) Presentation: This phase includes presentation of all exercises of importance for learning to read. New material is introduced, activities are demonstrated and the teacher checks pupil understanding of the new material. c) Guided practice: Pupils practice with the material under the guidance of the teacher. d) Independent practice: Pupils are given the opportunity to independently apply what has been learned; the teacher provides feedback and corrects pupils as needed. e) Review after a week. f) Review after a month. In DI, the teacher plays a highly influential role and both the process and the results are unambiguous.

In other research, Rosenshine, Meister and Chapman (1996) noted the importance of the aforementioned teaching functions for helping learners perform independently on highly structured tasks such as computational skills. 'Teaching in small steps' was very important along with 'guiding pupil practice.' In addition, 'extensive practice' and organizational factors were associated with positive pupil learning outcomes.

When it comes to beginning reading, 'explicit instruction' is more effective than indirect teaching methods particularly for disadvantaged children (Bennet, Jordan, Long, & Wade, 1976; NICHD, 2000; Raudenbush, 2009; Rosenshine, 1979). Research shows dramatic reductions in the incidence of reading failure when explicit instruction is provided by the classroom teacher. However, this research begs the question of whether DI is the most effective instructional approach for all children and particularly those children who have already made considerable reading progress.

Guided co-construction. According to Brown and Palincsar (1989), learning is the result of what can be called the processes of co-elaboration and co-construction. Both teachers and pupils are viewed as active participants in the construction of knowledge with ideas and experiences contributed by both as well (Mercer, 1995; Wells, 1999). Central to the guided co-construction and scaffolding of knowledge is the teacher talking with pupils in whole-class, group and individual contexts in order to guide their thinking.

In the domain of mathematics education, Freudenthal (1991) strongly opposed the presentation of mathematics as a formal system without a meaningful context and was thus a proponent of 'guided reinvention' or the generation of knowledge as opposed to the provision of knowledge (Rosenshine, Meister, & Chapman, 1996). Against this background, the instructional approach of 'guided co-construction' was designed and tested in a series of studies of the teaching of mathematics in primary education (Terwel, van Oers, van Dijk, & van den Eeden, 2009; van Dijk, van Oers, & Terwel, 2003). GCC proved not only feasible in real classroom settings but also effective in terms of learning gains when compared to a control group in which mathematics was directly instructed. The question which remains, however, is whether GCC can be successfully adapted and implemented for beginning reading instruction.

The instructional approach of GCC entails the following three core elements.

1. 'Guided' refers to the explicit role of the teacher for whole-class instruction and the scaffolding of pupils either in groups or individually.
2. 'Co-' refers to cooperative learning as an essential component of the use of reading as a cultural tool.
3. 'Construction' refers to the recognition and construction of symbols, words, sentences and so forth by pupils on the basis of their prior knowledge and experiences.

Taken together, these elements imply that teachers can facilitate beginning reading by presenting graphemes, phonemes, words and sentences but also elicit and scaffold contributions and constructions from pupils within a meaningful context. In this interactive

process, the differences between pupils are actually called upon; the phonics repertoire of letters and words is not only prescribed ahead of time but also created by the pupils as they interact and move along. And such a process is often called co-elaboration, co-construction or the guided reinvention of the language symbol system (Brown & Palincsar, 1989; Dewey, 1943; Freudenthal, 1991).

A question to be answered, however, is whether GCC can be used with success for early reading instruction. DI has shown itself to be effective for teaching children to read and particularly children with lower prerequisite skills. What about GCC?

2.1.4 Research question and specific hypothesis

DI has been shown to be an effective teaching approach in many domains and contexts (de Jager, 2002). Significant effects of DI have also been demonstrated in several beginning reading studies. However, we do not know if all children equally benefit from DI. In a recent meta-analysis, Slavin et al. (2009) found strong evidence for the effectiveness of several beginning reading programmes with cooperative learning approaches at their core. The relevant studies included schools with pupils from both higher and lower socio-cultural backgrounds.

In the present research, an intervention study was therefore designed to compare a DI approach to a GCC approach for the teaching of early reading skills. The research question was whether it is better for beginning reading instruction to provide pupils with letter-sound relations and ready-made words (i.e., DI) or scaffold pupil learning by helping them analyze and generate their own letter-sound relations and words in cooperation with both peers and teachers (i.e., GCC)? In addition to this, it was also asked if pupils from minority versus majority socio-cultural backgrounds might benefit differentially from the two instructional approaches. There are indications, for example, that minority pupils may benefit less from instructional approaches which require considerable verbal interaction, such as GCC, less than majority children do.

Based on a series of research projects (Terwel, van Oers, van Dijk, & van den Eeden, 2009; van Dijk, van Oers, & Terwel, 2003;) it was hypothesized that the word recognition skills of first grade children who received GCC would exceed the word recognition skills of first grade children who received DI. It was also hypothesized that the difference would be found for all measures of word recognition in the first grade. It was further hypothesized that the socio-cultural background of pupils would differentially affect their reading development: Minority pupils could benefit more - than majority pupils - from direct instruction (DI) and may profit less from teaching approaches which rely upon verbal interaction and initiative taking (GCC) There are some indications from literature that differences in home education play a major role in the differences observed among the pupils from different socio-cultural backgrounds (Leseman & de Jong, 1998).

2.2 Methods

2.2.1 Research Design and Participants

For this field experiment, a quasi-experimental pretest-posttest control group design was adopted. The experimental group or GCC group consisted of four classes with a total of 88 pupils. The control group or DI group consisted of five classes with 90 pupils. The participating schools were located in or near the Dutch city of Utrecht and taught their first grade pupils using the standard Dutch beginning reading and spelling programme Learning to Read Safely (see below). After intake interviews, the schools were classified in such a manner that different types of schools were represented across the experimental and control conditions. In such a manner, schools with predominantly minority pupils were equally distributed across the experimental and control conditions, just as schools with predominantly majority pupils from rural areas around the city of Utrecht. The experiment took place in real classroom situations. Random assignment of pupils, teachers and classes was not possible. However, after carefully assigning classes to treatments it turned out that no significant differences between the conditions were found on all pre-reading measures. In order to take these non-significant differences into account, all prereading measures were included in the analyses. All of the schools had the same denomination, namely Catholic.

The average age of the participants at the time of initial testing was 6 years and 4 months (SD = 5.1 months). Of the 178 pupils included in the study, 91 were male (51%) and 87 were female (49%). The socio-cultural backgrounds of the pupils were determined using data provided by the school administrations. Majority (i.e., native Dutch) pupils were identified (n=109 or 61%) and minority pupils — most of whom had Turkish or Moroccan backgrounds (n=69 or 39%). Of the 109 majority pupils, 56 were in the experimental group and 53 in the control group. Of the 69 minority pupils, 32 were in the experimental group and 37 the control group.

The socioeconomic status (SES) of the pupils was determined on the basis of parental education: 11 majority pupils and 43 minority pupils had two parents with a lower education, which was defined as low SES; 98 majority pupils and 26 minority pupils had one or two parents with a higher education, which was defined as high SES. In others words, most pupils with lower educated parents were in the minority group and vice versa for the majority group.

2.2.2 Reading programme used in both conditions

In the Netherlands, the most frequently-used beginning reading programme is Veilig Leren Lezen (Learning to Read Safely) by Mommers, Verhoeven, Koekebacker, van der Linden, Stegeman and Warnaar (2003). Two periods are distinguished in this reading programme:

one for the first half of first grade and one for the second half. During the first half, letter-sound relationships stand central. This period encompasses both the partial and full alphabetic subphases of the second phase in Ehri's model. The teacher instructs the children on the identities of letters and their sounds with the presentation of sight words (e.g., m is for /m/ as in maan [moon]). In such a manner, children learn that words consist of graphemes and that each grapheme represents a specific sound (i.e., phoneme). The three graphemes in maan are pointed out (i.e., m-aa-n) and then it is pointed out that the individual graphemes represent the individual phonemes /m/-/aa/-/n/ which can be merged to pronounce the word /maan/. The children also learn that meaning must be assigned to the word /maan/. The three steps in 'the fundamental reading operation' are also taught as part of the Learning to Read Safely programme: 1) linking graphemes to phonemes from left to right and thus in the direction of reading, which entails the visual analysis of graphemes, linking of phonemes to graphemes and remembering phonemes in sequence; 2) auditory synthesis or the merging of phonemes; and 3) the assignment of meaning.

In the second half of first grade, the automatization of word recognition stands central, the pupils are taught to read texts fluently and thus the consolidated alphabetic phase in Ehri's model (Aarnoutse, Beernink, & Verhagen, 2010; Verhagen, Aarnoutse, & van Leeuwe, 2006). It is important that the reading process become increasingly automated. This concerns not only the links between graphemes and phonemes but also the links between letter clusters and syllables.

In the present study, the Learning to Read Safely programme was used in both conditions but implemented differently by the teachers, as described in the following section. Prior to and during this study, the first author intensively guided the teachers in the experimental and control conditions. Prior to the start of each lesson, the teachers were given a teacher guide. The principles of the programme, the reading exercises, the role of the teacher and the role of the pupils were explained and discussed in great detail. If necessary, teaching activities were also demonstrated. The first author then visited the teachers in the control group and the experimental group every six weeks during the course of the present investigation to answer any questions about the programme, deliver the teacher guides for the upcoming units, again discuss the role of the teacher and the pupils and, finally, monitor just how well the programme was being implemented.

2.2.3 Characteristics of the two instructional approaches used in the classrooms

The Learning to Read Safely programme for grade 1 reading instruction was designed to be implemented step-by-step, which constitutes a form of Direct Instruction (DI). On the basis of the same Learning to Read Safely programme, an experimental teaching-coaching approach was developed, which constitutes a form of Guided Co-Construction (GCC).

In the two conditions, teacher training on different instructional principles was provided

and additional materials were developed and supplied to facilitate either DI or GCC. The teachers who used DI, for example, introduced new material for the pupils to practice in a demonstrated step-by-step manner. To illustrate, in a DI group, the graphemes r-v-i-s-p-aa-e were hung out on a string. In the introduction to a lesson, the teacher reviewed the graphemes covered in the previous lesson by reciting them. The teacher next introduced a new grapheme and pronounced the associated phoneme while showing the relevant grapheme card. The card was then added to the string. Next, the teacher asked the pupils if they knew of any words which began with the grapheme which was just being learned. The words were written on the blackboard with the grapheme written in a contrasting colour. Finally, the pupils worked individually in their workbooks on exercises in which the new grapheme stands central. This example shows control of the learning activities to be in the hands of the teacher. The teacher decides 'what' activity will be done and 'how' the pupils should do it.

In contrast, the teachers who used GCC introduced new material but gave the pupils an opportunity to exchange their knowledge of the new material and experiences with it (i.e., peer collaboration). In such a manner, pupils were allowed to construct their knowledge of the material right from the start, learn from each other and possibly learn more than just the presented material. The lesson described above for a DI group thus looked very different in the GCC group. All of the graphemes to be presented during the year, moreover, were hanging in the classroom. Only a sheet of paper hung between the graphemes already taught and the graphemes still to be taught. Similar to the DI group, old material was reviewed and new material was introduced at the beginning of the lesson. Thereafter, the pupils in the GCC group were given time to practice with the new material — in this case a new letter — but also invited to try to read those letters which had yet to be taught. And, somewhat different than in the DI group, they were invited to mention words which begin with that letter. The suggested words were written on the blackboard with the target grapheme in a contrasting colour and, in such a manner, the pupils practiced with not only the grapheme/phoneme mappings from the lesson but also the other grapheme/phoneme pairings mentioned by the pupils. More detailed information about the behaviour of teachers and pupils in both conditions will be discussed in a future article.

2.2.4 Measures

Tests of phonemic synthesis, letter knowledge, naming speed and phonemic analysis were administered in kindergarten to determine whether the experimental and control groups were equal with regard to the precursors to reading and, more specifically, their initial word recognition skill. Word recognition was subsequently tested on four occasions throughout the first grade.

All of the measures used in this study were administered in the schools by teachers-in-training who were also specially trained for this purpose. In several training sessions, the tests were practiced and their manuals discussed.

Phonemic Synthesis. This test measures the ability to reconstruct a word from its constituent phonemes (Aarnoutse & Verhagen, 2001). The 20 items range in difficulty from words like ijs (ice) to words like paraplu (umbrella). The Cronbach's α in the Aarnoutse and Verhagen study was .89.

Letter Knowledge. A test developed by Aarnoutse, Beernink and Verhagen (2010) was used to measure the children's passive letter knowledge. The test consists of 23 lists of 23 letters each with x, y and q not included and two of the 23 letters, the s and o, serving as practice items. For each list, a single letter is read aloud and the child is asked to circle the letter which has been read aloud. The Cronbach's α in the Aarnoutse, Beernink and Verhagen study was .92.

Naming Speed for Letters/Digits. In each of these tests, as developed by Aarnoutse, Beernink and Verhagen (2010), five columns of 10 items each are presented; the first column is a practice column. The child is asked to name the items in the columns as quickly but accurately as possible. The child's score is the time required in seconds to name the 40 items. Naming Speed for Letters uses the letters o, s, m, p and k because these letters are most familiar to kindergarten children. The test-retest reliability mentioned in the manual is .88. Naming Speed for Digits uses the numbers 1, 2, 3, 4 and 5. The test-retest reliability mentioned in the manual is .86.

Phonemic Analysis. This test measures the child's ability to analyze a pseudoword into its constituent phonemes (Verhagen & Aarnoutse, 2001). The child is asked to listen to a series of 40 pseudowords and name the first phoneme words like buin and krontebel on 20 occasions and name the last phoneme in words like koes and draap on 20 occasions. A Cronbach's α of .94 has been reported by Verhagen and Aarnoutse (2001).

Speed of Word Recognition. This is a measure of the child's ability to decode printed words (Aarnoutse & Kapinga, 2007). The child is presented a card with a list of 100 words of increasing difficulty. The unrelated words range from simply words like raam [window] to multi-syllabic words like trekdier [draught animal]. The child is asked to read the words aloud as quickly as he or she can but without pressure. The test score is the number of words read correctly in 90 seconds. The test has two parallel forms. The test-retest correlations mentioned in the manual exceed .86.

The Speed of Word Recognition test was administered in November, January, March and May of the first grade.

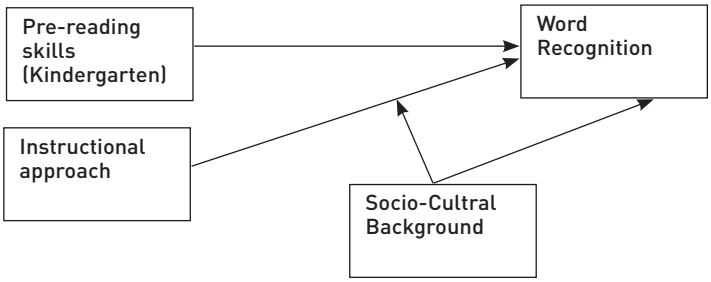
2.2.5 Data analyses

Given that most pupils cannot read at the start of grade 1, the Word Recognition test could not be administered as a pre-test. We therefore used the kindergarten tests Phonemic

Synthesis, Letter Knowledge, Naming Speed for Letters/Digits and Phonemic Analysis as the pre-tests. All of these tests are known to be important predictors of later word recognition (Aarnoutse, 2004; Aarnoutse, van Leeuwe, & Verhoeven, 2000, 2005; Beernink, 2002; Verhagen, Aarnoutse, & van Leeuwe, 2006, 2008).

In Figure 1, the conceptual model underlying this study is presented. The two longest arrows indicate direct effects of pre-reading skills and grade 1 reading instruction on the children's grade 1 word recognition. The two shorter arrows - originating from Socio-cultural background - represent the direct and the moderator effect of socio-cultural background on the children's grade 1 word recognition. A moderator variable has an impact on the relation between Reading Instruction and Word Recognition (vgl. Holmbeck, 1997). In figure 1 it is hypothesized that pupils from different Socio-Cultural backgrounds differentially benefit from Reading Instruction.

Figure 1: Conceptual model guiding analyses.



Whether or not the grade 1 word recognition of the children instructed using GCC exceeds the grade 1 word recognition of the children instructed using DI was analyzed by first determining if the two groups differed with respect to gender and socio-cultural background using chi-square tests. Then whether they differed with respect to kindergarten literacy using t-tests; t-tests were then applied to test for significant differences in the word recognition of the experimental versus control groups across grade 1.

Whether or not the children from minority versus majority socio-cultural backgrounds benefitted differently from the different types of instruction was analyzed by examining the effects of Instructional Programme (i.e., condition) on their grade 1 word recognition after controlling for gender, socio-cultural background and kindergarten literacy. First, a fixed-effects model which included the interaction between gender and condition, on the one hand, and the interaction between socio-cultural background and condition, on the other hand, was tested. Second, all non-significant interactions were removed and the model was retested. In the next step, the non-significant main effects were dropped with

the exception of the main condition effect even when it was not significant. The model determined in such a manner was then referred to as the final model.

In the final set of analyses, possible differences in the development of word recognition skills were investigated in repeated measures analyses of variance for the majority versus minority pupils in the control versus experimental groups.

2.3 Results

In Table 1, the means and standard deviations for the tests administered in kindergarten and grade 1 are presented for the experimental and control groups separately and for the experimental and control groups according to socio-cultural background. The Word Recognition (WR) means can be seen to increase over time for all of the groups although the scores in the experimental group are generally higher than the scores in the control group. Within the experimental group, moreover, the majority pupils score higher than the minority pupils. But within the control group, the opposite is found: The minority pupils score higher than the majority pupils.

The percentages of boys and girls in the control versus experimental groups did not differ significantly [Chi-square = is .092, df = 1, p = .762]. Similarly, the percentages of majority versus minority pupils in the control versus experimental groups did not differ significantly [Chi-square = .422, df = 1, p = .516].

T-tests performed under the assumption of unequal variances showed the control versus experimental groups do not differ on any pre-reading skills (see Table 2).

In Table 3, the mean WR scores on four occasions throughout grade 1 consistently showed the experimental group to outperform the control group. The t-tests, however, showed only statistically significant differences on Word Recognition 1 and 2 (i.e., the first two measurement occasions).

Table 1: Mean Test Scores (Standard Deviations) for Control and Experimental Groups, also according to Socio-Cultural Backgrounds of Students.

	Total (N=178)	Experimental (N=88)	Control (N=90)	Exp. min (N=32)	Exp. maj. (N=56)	Con. min (N=37)	Con. maj (N=53)
Phonemic Synthesis	12.07 (5.45)	12.76 (5.02)	11.39 (5.79)	11.88 (5.75)	13.27 (4.53)	11.46 (6.05)	11.34 (5.66)
Letter Knowledge	12.24 (6.21)	12.06 (6.57)	12.41 (5.86)	12.66 (6.33)	11.71 (6.74)	12.05 (5.55)	12.66 (6.10)
Phonemic Analysis	32.43 (8.35)	32.56 (8.82)	32.31 (7.91)	31.78 (10.35)	33.00 (7.89)	31.68 (7.83)	32.75 (8.00)
Naming Speed Digits	44.76 (13.96)	44.09 (15.13)	45.42 (12.75)	42.10 (13.76)	45.23 (15.87)	44.75 (11.49)	45.89 (13.65)
Naming Speed Letters	59.21 (43.67)	62.65 (53.78)	55.84 (30.71)	67.53 (78.69)	59.86 (32.60)	55.94 (25.67)	55.77 (34.03)
Word Recognition test 1	28.03 (16.89)	32.45 (20.14)	23.71 (11.52)	31.94 (16.62)	32.75 (22.04)	24.89 (8.96)	22.89 (13.03)
Word Recognition test 2	39.52 (19.59)	42.72 (22.02)	36.39 (16.40)	39.84 (19.80)	44.36 (23.21)	38.62 (12.68)	34.83 (18.51)
Word Recognition test 3	54.51 (19.75)	56.89 (21.59)	52.18 (17.58)	55.09 (17.64)	57.91 (23.64)	57.97 (13.60)	48.13 (18.98)
Word Recognition test 4	63.65 (20.70)	65.03 (22.09)	62.29 (19.27)	62.13 (18.40)	66.70 (23.94)	70.54 (14.09)	56.53 (20.39)

Table 2: Tests for Differences between Control (N=90) and Experimental (N=88) Groups on Kindergarten Pre-reading Measures.

	Group	Mean	SD	t	df	p
Phonemic Synthesis	Control	11.39	5.79	-1.69	73.53	.09
	Experimental	12.76	5.02			
Letter Knowledge	Control	12.41	5.86	.38	172.76	71
	Experimental	12.06	6.57			
Phonemic Analysis	Control	32.31	7.91	-.20	172.99	.85
	Experimental	32.56	8.82			
Naming Speed Digits	Control	45.42	12.75	.63	169.76	.53
	Experimental	44.09	15.13			
Naming Speed Letters	Control	55.84	30.71	-1.03	137.65	.30
	Experimental	62.65	53.78			

Table 3: Test for Differences between Control and Experimental Groups on Word Recognition measured on Four Occasions in First Grade.

	Group	Mean	SD	t	df	p
Word Recognition 1	Control	23.71	11.52	-3.54	137.76	.00
	Experimental	32.45	20.14			
Word Recognition 2	Control	36.39	16.40	-2.17	160.71	.03
	Experimental	42.72	22.02			
Word Recognition 3	Control	52.18	17.58	-1.59	167.52	.11
	Experimental	56.89	21.59			
Word Recognition 4	Control	62.29	19.27	-.88	171.70	.38
	Experimental	65.03	22.09			

In Table 4, the results of an ANOVA with the initial measurement of WR in November of grade 1 as the dependent variable can be seen to show the kindergarten variables of Phonemic Synthesis, Letter Knowledge and Naming Speed for Digits but not Phonemic Analysis or Naming Speed for Letters to play a significant role in the children's early WR with the experimental group performing better than the control group.

Table 4: ANOVA Results for Final Model of Word Recognition 1.

Source	Type III Sum of Squares	df	Mean Square	F	p	Partial Eta Squared
Corrected Model	24925.05	4	6231.26	42.12	.00	.49
Intercept	6568.17	1	6568.17	44.40	.00	.20
Condition	2396.78	1	2396.78	16.20	.00	.09
Phonemic Synthesis	1581.48	1	1581.48	10.69	.00	.06
Letter Knowledge	1508.54	1	1508.54	10.20	.00	.06
Naming Speed Digits	5896.07	1	5896.07	39.86	.00	.19
Error	25592.75	173	147.94			
Total	190406.00	178				
Corrected Total	50517.80	177				

In Table 5, the ANOVA results are summarized for WR in January of grade 1 (i.e., measurement occasion 2). In addition to a treatment effect in favour of the experimental group in the final model, the kindergarten variables of Phonemic Synthesis, Letter Knowledge and Naming Speed for Digits are again found to play a significant role in the children's WR. The difference between the two conditions was less on occasion 2 than on occasion 1, however.

In Table 6, the ANOVA results are summarized for WR in March of grade 1 (i.e., measurement occasion 3). The results are very different than on the previous measurement occasions. The significant main effect of condition is no longer found but, instead, a significant interaction between condition and the socio-cultural backgrounds of the pupils: Pupils from different social backgrounds learn differently from different types of instruction. The estimated WR means for the majority versus minority pupils in the experimental group were 56.41 and 52.93, respectively. For the control group, the estimated means were 52.28 and 58.10, respectively. The majority pupils thus score better than the minority pupils in the

Table 5: ANOVA Results for Final Model of Word Recognition 2.

Source	Type III Sum of Squares	df	Mean Square	F	p	Partial Eta Squared
Corrected Model	32411.84	4	8102.96	39.50	.00	.48
Intercept	13838.89	1	13838.89	67.45	.00	.28
Condition	1055.40	1	1055.40	5.14	.03	.03
Phonemic Synthesis	1464.82	1	1464.82	7.14	.01	.04
Letter Knowledge	2261.72	1	2261.72	11.02	.00	.06
Naming Speed Digits	9368.14	1	9368.14	45.66	.00	.21
Error	35492.61	173	205.16			
Total	345866.00	178				
Corrected Total	67904.45	177				

Table 6: ANOVA Results for Final Model of Word Recognition 3.

Source	Type III Sum of Squares	df	Mean Square	F	p	Partial Eta Squared
Corrected Model	31692.28	5	6338.46	29.19	.00	.46
Intercept	34728.77	1	34728.77	159.92	.00	.48
Condition	456.02	1	456.02	2.10	.15	.01
Letter Knowledge	3117.56	1	3117.56	14.36	.00	.08
Naming Speed Digits	11753.58	1	11753.58	54.12	.00	.24
Condition* background	2635.87	2	1317.93	6.07	.00	.07
Error	37352.22	172	217.16			
Total	597858.00	178				
Corrected Total	69044.49	177				

experimental group while the minority pupils score better than the majority pupils in the control group. Naming Speed for Digits in kindergarten is again found to be an important predictor of WR; kindergarten Letter Knowledge is also important but to a lesser extent than Naming Speed for Digits, as also found on previous WR measurement occasions.

The results in Table 7 for the measurement of WR in May of grade 1 (i.e., the fourth measurement occasion) again show a significant interaction between condition and socio-cultural background and also a significant effect of condition at the 5% level but now in favour of the control group. The estimated WR means for the majority versus minority pupils in the experimental group are 64.71 versus 60.96; in the control group, they are 62.19 versus 70.24 . Once again, thus, the majority pupils score better in the experimental group but the **minority** pupils score better in the control group. The results in Table 7 show Naming Speed for Digits to again be an important predictor of WR. The contribution of Naming Speed for Letters to WR is now significant as well.

Table 7: ANOVA Results for Final Model of Word Recognition 4.

Source	Type III Sum of Squares	df	Mean Square	F	p	Partial Eta Squared
Corrected Model	30.835.31	4	6167.06	25.58	.00	.41
Intercept	152093.04	1	152093.04	581.60	.00	.77
Condition	1454.62	1	1454.62	5.56	.02	.03
Naming Speed Digits	15566.11	1	15566.11	59.52	.00	.26
Naming Speed Letters	2419.30	1	2419.30	9.25	.00	.05
Condition *background	4563.02	2	2281.51	8.72	.00	.09
Error	44979.39	173	261.51			
Total	796861.00	178				
Corrected Total	75814.70	177				

From the foregoing preliminary analyses as presented in table 4-7, some tentative conclusions were drawn from the cross sectional analyses of variance. Now we turn to the final test of the hypothesis in a longitudinal, repeated measures analysis.

To obtain a parsimonious but realistic model of the development of WR under different instructional conditions, we decided to enter all of the effects which were found to be significant in one of the preliminary analyses into a repeated measures analysis of WR over time. The following were thus included: main effects of kindergarten Phonemic Synthesis, Letter knowledge, Naming Speed for Digits and Naming Speed for Letters; the interaction between condition and socio-cultural background; and the main effect of condition. Given that socio-cultural background was a diversification variable and not a covariable as in the previous analyses, it was decided to include socio-cultural background as an independent variable in the repeated measures analysis. Non-significant interactions and main effects were next removed successively from the model. The final results for the repeated measure of WR1 through WR4 over time are presented in Table 8. The within subjects test statistics were calculated using multivariate Wilks' lambda F-tests. The between subjects test statistics were derived from type III sum of squares.

Table 8: Results of Repeated Measures Analysis of Word Recognition over Time (WR1 to WR4).

Effect	F	df1	df2	p	part.eta.sq.
Within subjects					
Time	22.87	3	169	.00	.29
Time*Condition	4.01	3	169	.01	.07
Time*Naming Speed Letters	0.40	3	169	.76	.01
Time*Letter Knowledge	3.42	3	169	.02	.06
Time*Naming Speed Digits	2.16	3	169	.10	.04
Time*Condition* Background	3.83	6	338	.00	.06
Between subjects					
Condition	4.22	1	171	.04	.02
Naming Speed Letters	5.75	1	171	.02	.03
Letter Knowledge	9.62	1	171	.00	.05
Naming Speed Digits	54.01	1	171	.00	.24
Condition*Background	4.86	2	171	.01	.05

The last column in Table 8 shows Naming Speed for Digits to be by far the best predictor of WR. The main effect of Condition was significant at the 5% level in favour of the experimental group. The Time by Condition interaction was significant, which shows that the developmental patterns for WR1 to WR4 are not parallel for the experimental versus control groups. The significant Time by Condition by Socio-cultural background interaction similarly shows the development of the minority versus majority pupils in the different conditions are not parallel over time. The effects reported here, when the influence of possibly confounding variables has been taken into consideration, are presented in table 9 and visualized in Figure 2.

Table 9: Estimated means of WR1 to WR4 by control and experimental group.

	WR1	WR2	WR3	WR4
Control group	23.91	36.79	53.11	63.61
Experimental group	32.06	41.70	56.12	64.03

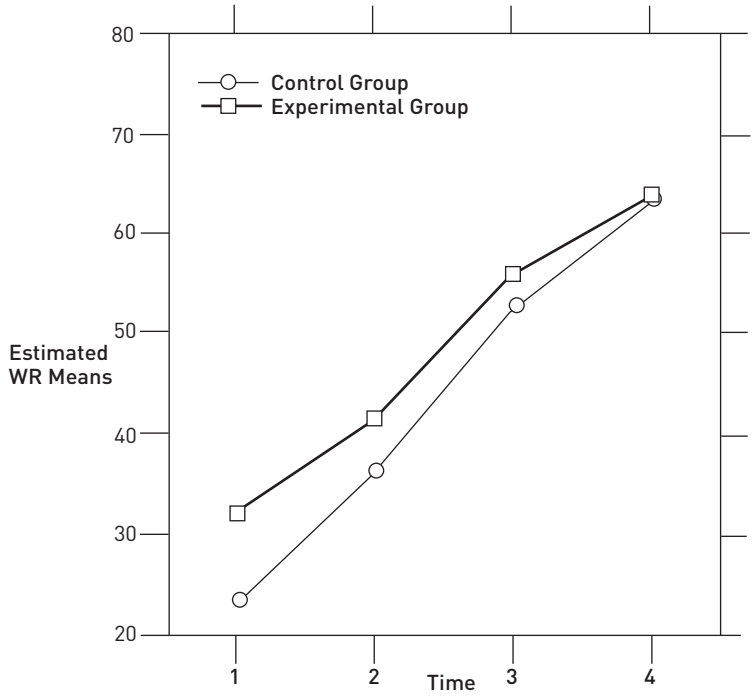


Figure 2: Development of word recognition in control versus experimental groups.

In figure 3 the results are depicted by four separate lines for each of the four student categories: outcomes on Word Recognition are broken down by intervention and socio-cultural background.

From the estimated means presented in table 9 and the corresponding graph's in figure 2 we conclude that the learning gains of pupils in both conditions are considerably. Looking at the results over the whole year, the overall conclusion from the repeated measures analysis in table 8 is substantiated by the marginal means presented in table 9. Pupils in the Guided Co-Construction (experimental) condition outperform their counterparts in the control group. However, the differences between the estimated means in both conditions gradually diminish. In the final quarter of the school year, there appears to be a process of levelling off in both conditions. However, the curve of the experimental group seems to flatten most in the experimental group. At the last measurement the pupils in the Direct Instruction (control) group almost catch up to the level of the experimental group.

From the repeated measurements in Table 8 it was concluded that besides the main effect in favour of the Guide Co-construction condition, interaction effects occur. In our research we are especially interested in differential effects of condition by social background. Do pupils from various social background differentially benefit from Guided Co-construction? The results are presented in table 10 and visualized in Figure 3.

Table 10: Estimated means of WR1 to WR4 by control and experimental group and by socio-cultural background.

	WR1	WR2	WR3	WR4
Control group minorty	24.85	38.59	57.85	70.34
Control group majority	22.97	35.00	48.37	56.89
Experimental group minorty	30.76	38.30	53.68	60.85
Experimental group majority	33.37	45.10	58.57	67.22

Inspection of table 10 and Figure 3 makes clear that the minority pupils in the control group and majority pupils in the experimental group end up having the highest means. This shows the minority pupils to benefit most from DI and the majority pupils to benefit most from GCC. The minority pupils in the experimental group, in particular, make less progress than all of the other pupils. If we look towards the 109 majority pupils, the performance differences between the conditions will remain more or less constant throughout the year.

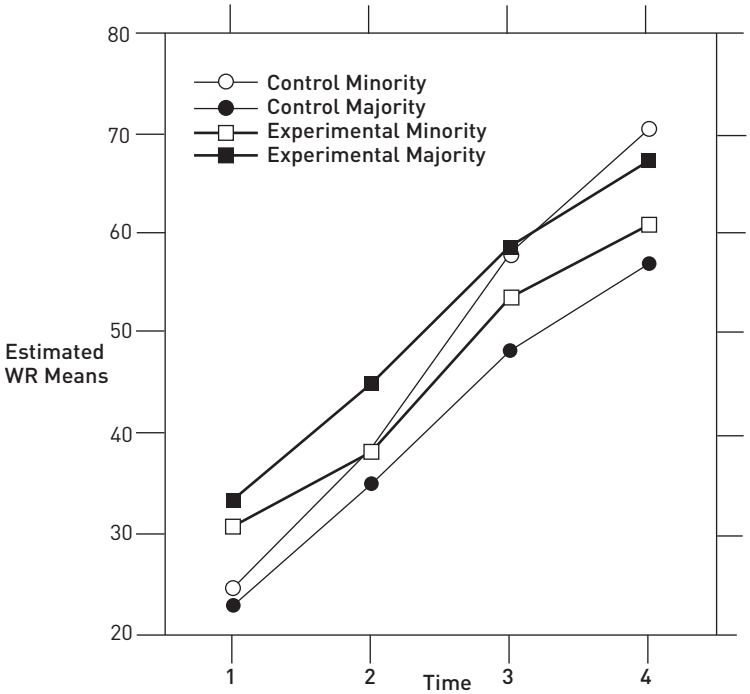


Figure 3: Development of word recognition according to socio-cultural backgrounds of students in control versus experimental groups

2.4 Conclusions and Discussion

In the present study, the effects of an experimental beginning reading programme were examined using a quasi-experimental pretest-posttest control group research design. The instructional approach used in the experimental group was guided co-construction (GCC). The instructional approach used in the control group was direct instruction (DI).

2.4.1 Conclusions

The results of a repeated measures analysis of the development of word recognition during the initial stages of learning to read (i.e., throughout first grade) showed a main effect of condition, with pupils in the experimental group outperforming pupils in the control group. Our hypothesis - that the word recognition skills of first grade children who received a GCC approach to beginning reading instruction would exceed the word recognition skills of first grade children who received DI - appears to be confirmed in the final, most comprehen-

sive analysis using repeated measurement. However, significant interactions and closer inspection of the children's WR across the year showed clearly different effects over time: a significant condition effect in favour of GCC on WR1 and WR2; a nonsignificant condition effect on WR3; and a significant condition effect in favour of DI on WR4. In other words, the positive effects of GCC compared to DI disappeared by the end of first grade. The better performance by the experimental group attenuated over time with better performance by the control group on the last measurement. DI was found to be more effective than GCC.

In addition to a main effect of condition, a significant interaction between condition and the socio-cultural background of the pupils was found. Closer inspection of the results showed marked variation over time: no significant interaction between condition and socio-cultural background during the first half of the year; however, during the second half of the year, the majority pupils in the GCC group (N=56) scored better than the minority pupils in this group (N=32) and the minority pupils in the DI group (N=37) scored better than the majority pupils in the group (N=53). With the exception of the minority pupils in the control group, the WR of all of the subgroups also developed in parallel. Naming Speed for Digits was found to be the best predictor of WR, but this finding will be discussed in a future article.

2.4.2 Discussion

Before discussing the findings of this study, some possible limitations should be mentioned. Firstly, The number of minority pupils in the two instructional conditions was small (32/37) and the standard deviations were large. There was thus considerable variability in the children's performance. Whether or not the findings in this study hold for pupils at other schools is therefore open to question. And secondly, we used kindergarten tests as the pre-tests (all these tests are known to be important predictors of later word recognition). Despite the fact that no significant differences on these pretests between the experimental and the control group were found, the possibility exists that the experimental group was initially in a better position to acquire reading skills i.e. word recognition. We could not use a Word Recognition pre-test, given that most pupils cannot read at the start of grade 1. The first test in Word Recognition was assigned two and a half months after the start of the experiment as the first effect measure in a series of four measurements. Given the fact that random assignment was not possible, it can not be excluded that there were differences between the experimental and control group. The difference between both groups on the first Word Recognition test was rather large and disappearing in the course of the experiment. This could mean that the effects can not be fully attributed to the treatment. Therefore we recommend further research.

Our findings are nevertheless in line with Raudenbush' recommendation of 'explicit instruction' for disadvantaged children in particular (2009). In general, the present findings also resemble the findings of a meta-analysis recently conducted by Slavin et al. (2009); strong evidence was provided for the effectiveness of beginning reading instruction which

has cooperative learning at its core. Although GCC has a cooperative learning component, a one-to-one parallel to the instructional methods included in the meta-analysis by Slavin et al. does not exist.

In order to explain the present findings, we must broaden our perspective and examine other studies conducted in other domains with other age groups. The present findings are in line with earlier findings from a series of studies in the domain of primary mathematics. GCC proved to be particularly effective when compared to a 'providing' instructional approach. Similarly, in the present study, explicit domain-specific instruction and scaffolding of the co-construction and co-elaboration of the beginning reading process was found to be pivotal. The elicitation and elaboration of the letter, sound, word and sentence knowledge of pupils and sharing of this knowledge in a collaborative manner constituted a particularly effective instructional approach. Connecting new and existing knowledge to the experiences of children via the use of meaningful contexts is considered a critical aspect of GCC. Just as for the 'language' of mathematics, children can thus be guided to decontextualize and recontextualize their reading knowledge. However, GCC only proved more effective than DI during the first half of first grade when 'the fundamental reading operation' stands central. In the final quarter of the year, when the 'automatization of word recognition' stands central, DI was found to be more effective than GCC. It thus seems likely that when it comes to speed and automatization of word recognition, a structured leading role for the teacher (DI) can be more effective than GCC. Where contributions from different pupils are called for, GCC appears to be fruitful.

The finding of a significant interaction between type of instruction and socio-cultural background but only during the second half of first grade is consistent with the findings of the Dutch Inspection of Education (2006e, 2008b) and Stoep (2008) who both show early reading performance to only relate to socio-economic status or ethnic background when pupils must learn to read longer, more complex words and at a faster rate later in their reading development. Droop and Verhoeven (2003), Verhoeven (2000) and Leseman and de Jong (1998, 2001) also report similar findings. In this later period of reading development, minority pupils appear to benefit more from DI and majority pupils more from GCC. The significant interaction between type of instruction and socio-cultural background only towards the end of the year remains difficult to explain. One possible explanation can be found in differences between socio-cultural groups in 'home literacy', as described by Leseman and de Jong (1998, 2001). The home literacy practices of majority parents resembles GCC and thus places these pupils in a better position to benefit from GCC than minority pupils. Both minority and majority pupils initially benefit from GCC for the acquisition of knowledge but later, during the second half of first grade, this is no longer the case. The characteristics of DI resemble the home cultures and home literacy of children coming from minority families and lower SES backgrounds, and this resemblance explains why minority pupils later benefit more from DI than GCC. DI entails more centralized teacher instruction and guidance than GCC and also places a greater emphasis on precision than GCC. Teachers give directions to practice and explain more complicated phenomena. There is less cooperation between

pupils during DI and therefore less ambiguity, less reliance on prior experience and prior knowledge than during GCC. DI requires less initiative from the learner than GCC and does not emphasize the construction or sharing of knowledge with the teacher and other pupils while GCC does. In contrast to the majority pupils in the present study, the minority pupils, who have less experience with the verbal skills required to collaborate and interact during GCC, benefited most from DI and least from GCC.

In closing, it can be stated that first-grade pupils receiving GCC generally outperformed first-grade pupils receiving DI. This effect faded during the second half of first grade when minority pupils appeared to benefit more from a DI approach to their further reading instruction when compared to minority pupils receiving GCC.