Development of beginning reading

A study in word recognition during the first primary school year

ISBN/EAN: 978-90-8928-085-5

Print: Netzodruk
Cover illustration: Maaike Disco
Layout: Maaike Disco, Aletta de Boer

Vrije Universiteit

Development of beginning reading

A study in word recognition during the first primary school year

Academisch proefschrift

ter verkrijging van de graad van doctor aan
De Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Psychologie en Pedagogiek
op vrijdag 31 oktober 2014 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Maria Johanna Snel

geboren te Zaandam

Promotor: Prof. dr. H.J.M. van Oers

Copromotoren: Prof. dr. J. Terwel
Prof. dr. C.A.J. Aarnoutse
Dr. ing. W.M. van der Veld

Voorwoord

In de afgelopen jaren ben ik in de bevoorrechte positie geweest door een deel van mijn tijd te mogen werken aan een proefschrift over het leesonderwijs en de leesontwikkeling van de kinderen in groep 3. In deze periode is me duidelijk geworden dat je weliswaar alleen aan een proefschrift werkt, maar dat je er tegelijkertijd nooit echt alleen voor staat. Er zijn dan ook diverse mensen die ik zou willen bedanken voor het werk dat zij met en voor mij hebben gedaan tijdens dit intensieve traject.

Als eerste zou ik graag mijn (co)promotoren prof. dr. Jan Terwel, prof. dr. Cor Aarnoutse, dr. Jan van Leeuwe, dr. ing. William van der Veld en prof. dr. Bert van Oers willen bedanken. Jan (Terwel), je had het vertrouwen in mij en gaf mij deze promotieplek. Je bijdragen aan dit proefschrift waren onmisbaar. Door de vragen die je stelde liet je me nadenken om zo op de juiste manier de keuzes te kunnen verwoorden en verantwoorden. Ik heb je ervaren als een begeleider die ook de kleine details wilde doorgronden. Daarnaast heb ik onze persoonlijke gesprekken tijdens de vele autoritjes naar Nijmegen als zeer waardevol ervaren.

Cor, dankzij jou heeft dit proefschrift een gedegen vakinhoud gekregen, jouw bijdrage was hierin onmisbaar. Jouw expertise over het leesonderwijs hebben verschillende artikelen de juiste richting gegeven. Maar, ook je persoonlijke steun heb ik zeer gewaardeerd. Je mail-tjes en onze telefoongesprekken hebben me soms door lastige periodes heengetrokken.

Helaas, is Jan (van Leeuwe) tijdens mijn promotietraject overleden. Zijn bijdrage in met name het tweede hoofdstuk was van grote waarde. Samen hebben wij onze eerste analyses gedaan, waarbij steevast de vraag klonk “En Mirjam, wat betekenen deze resultaten nu?”.

William heeft daarna de werkzaamheden van Jan overgenomen. William, samen met jou heb ik veelvuldig gesproken over de betekenis van de resultaten voor de praktijk. Daarnaast kon ik met jou ook delen wat het betekent om onderzoek te doen, een baan te hebben gecombi-neerd met een gezin met jonge kinderen. Bedankt je voor je geduld, steun en humor.

Als laatste richt ik me tot prof. dr. Bert van Oers. Ondanks dat je rol beperkt was, is deze zeer waardevol geweest voor het afronden van dit proefschrift. Je hebt het voor mij moge-lijk gemaakt om binnen de gekozen setting het proces af te ronden. Bedankt voor je belang-stelling en je bijdrage in met name de laatste fase van het traject.

Dan wil ik hier zeker ook de leerkrachten noemen die een vitale bijdrage hebben geleverd aan het onderzoek op de werkvloer: Natascha Loffeld en Digna Verhagen (Mattheusschool, Utrecht), Marjan de Dood (Wijngaard, Doorn), Ans Elsinghorst en Gonny Wesselman (Lucas Batau, Nieuwegein), Elisabeth Sikkema (Mauritsschool, Nieuwegein), Judith Potman (St. Jozefschool, Oudewater), Bernadet Dankaart (Marcusschool, Utrecht), Paula Hom (De Twijn, Utrecht), Tetchen Steenbrink (Johannes XXIII, Utrecht). Ik wil jullie bedanken voor

jullie inzet, betrokkenheid en deskundigheid. Tevens wil ik de leerlingen bedanken voor het enthousiasme waarmee zij een bijdrage leverden aan het onderzoek. In de groepen op de basisscholen kwam dit onderzoek echt voor mij tot leven en kregen theoretische kaders praktische implicaties en uitwerkingen.

Vervolgens wil ik ook de Hogeschool Domstad en de Hogeschool Utrecht bedanken voor het vertrouwen en de tijd die ik heb gekregen om dit onderzoek uit te voeren. Tevens zou ik enkele studenten willen bedanken die mij geholpen hebben bij het verzamelen van de gegevens: Anne Blanksma, Linda Breeman, Claire Cox, Myrthe Drexhage, Carolijn Frankhuisen, Marleen Harmsel, Anne de Jong, Gerda de Jong, Esther Pantekoek, Jorinde Romanesko, Linda Schipper, Melanie Valentijn, Femke Wijnia, Tatjana Hansma, Sabrina Hunt, Annemiek Hooijer. En de twee oudste dochters van mijn man: Roos en Merel Klinkien.

Lieve familie en vrienden jullie wil ik in het bijzonder noemen. Als eerste mijn ouders, zij hebben mij in mijn jeugd een context gegeven waarin persoonlijke ontwikkeling gestimuleerd werd, zonder dat daarbij een norm werd opgelegd. Mijn moeder deed dat door de juiste voorwaarden te scheppen, mijn vader deed dat door het goede voorbeeld te geven. Ook na zijn eigen promotie bleef het verder verdiepen in zijn vak een geliefde bezigheid, een eigenschap die ik ook in mezelf herken. Daarnaast wil ik mijn paranimfen Marjolein Kool en Dominique Sluijsmans bedanken. Zij hebben mij vooral in de laatste fase bijzonder gesteund. Vele gesprekken van ons gingen over het doen van onderzoek, het werk op een Hogeschool en een goede balans tussen werk en privé. Jullie vriendschap is erg waardevol voor me.

Tenslotte richt ik me tot Theo, Mees en Anna. Theo, bedankt voor het vertrouwen dat je altijd in me hebt gehad. Maar vooral bedankt voor de voorwaarden die je geschapen hebt, waarin ik kon werken. Je steun hebben mij de rust gegeven om dit proefschrift te schrijven. Mees en Anna, jullie zijn tijdens het werken aan mijn proefschrift geboren. Jullie lieten me beseffen dat passie voor onderzoek en betrokkenheid op werk belangrijk is, maar gepaste afstand er de juiste zin en betekenis aan kunnen geven.

Mirjam

Table of contents

Chapter 1

General Introduction

1.1 Introduction

The importance of reading in our time is evident. Learning to read is important for both children and the future of society. If children do not learn to read, both their educational and professional possibilities are limited. Moreover, poor readers have been shown to have lower self-confidence and lower learning motivation than other children (Vernooy, 2007).

Most children successfully learn to read in primary school, but some 10% leave primary school as 'functionally illiterate' (Sijtstra, van der Schoot, & Hemker, 2002; Wentink & Verhoeven, 2001). This means that they do not have sufficient reading ability to function adequately in society. To reduce the percentage of functionally illiterate individuals in society, considerable work has been devoted to the detection and prevention of reading difficulties. For example, intermediate goals to be met for beginning reading and writing have been formulated for the Dutch language (Verhoeven & Aarnoutse, 1999). These goals indicate what schools should focus on in the teaching of reading and writing. In addition, considerable research has been conducted on the development of children's word recognition skills and reading instruction which is aimed at stimulating children's word recognition (Aarnoutse, van Leeuwe, Voeten, & Oud, 2001; Henneman, Kleijnen, & Smits, 2004; Houtveen, Brokamp, & Smits: 2012; Verhagen, 2009; Verhoeven & van Leeuwe, 2003; Wentink & Verhoeven, 2001, 2004; Wouters & Wentink, 2005). Word recognition can be defined in this respect as the process of converting a sequence of letters into sounds for the identification of a word (Coltheart, 1978).

To date, two conclusions can be drawn about the development and teaching of word recognition. First, prevention is better than cure. This means that good reading instruction should begin early (i.e., in kindergarten). With preparatory word recognition activities, those factors that are known to influence the early development of children's word recognition should be targeted and practiced: phonemic awareness, letter knowledge, and naming speed (Aarnoutse, 2004; Aarnoutse, van Leeuwe, & Verhoeven, 2000, 2005; Beernink, 2002; Verhagen, Aarnoutse, & van Leeuwe, 2006, 2008). Phonemic awareness is the ability to hear, identify and manipulate phonemes or the smallest units of sounds which differentiate meaning. Letter knowledge is the familiarity with the names and sounds associated with printed letters, while naming speed is the ability to rapidly name a random sequence of otherwise well-known letters, digits, pictures of well-known objects or colours. Second, most reading difficulties can be avoided when detected quickly enough and therefore initiate early interventions to prevent further stagnation. This means that primary school teachers must know about the development of word recognition skills and the most effective strategies to stimulate these.

This dissertation is concerned with the following two topics: (1) the development of children's word recognition during grade 1 and (2) the most effective instruction to stimulate this development.

In the following, we will first review the phases of reading development in general, early grade 1 reading instruction and two instructional approaches: Direct Instruction (DI) versus Guided Co-Construction (GCC). An outline of the word recognition research presented in this dissertation will then be presented.

1.1.1 The development of early reading skills

According to Chall (1996) children proceed through predictable phases when beginning to read. These phases have many similarities with the phases of reading development in Ehri's model (1998, 2005) and the three phases in the development of literacy formulated by Verhoeven and Aarnoutse (1999). Verhoeven and Aarnoutse labelled these phases as: emerging, early and advanced literacy.

During the first phase, children between 0 and 4 years of age learn to speak and understand spoken native language. This provides the basis for their later literacy and is referred to as 'emerging literacy' because the children learn about books, print and writing at this time and thus long before they can actually read or write and long before they go to school.

During the second phase in the development of literacy from the beginning of kindergarten (around the age of 4) through the end of grade 1 (around the age of 7), children first discover letter-sound relationships and thus the alphabetic principle of correspondence between letters (i.e., graphemes) and speech sounds (i.e., phonemes). This is crucial for their early literacy development and subsequent instruction. After about five months of formal reading instruction in grade 1 (around January of grade 1 for the average primary school pupil) a rapid increase in the speed of word recognition occurs. This is also crucial for early literacy as it allows greater attention to be paid to sentence comprehension than to the decoding of words for recognition. Efficient and thus quick but accurate word recognition is critical for the further development of literacy.

During the third phase in the development of literacy, word recognition gets automatized. Children from grade 2 to grade 6 increasingly recognize words as a whole, which means that the parts of a word no longer have to be individually decoded to identify the word and its meaning within a particular context. Word recognition becomes largely automatic and therefore less effortful. And greater attention can be paid to text comprehension than to individual word and sentence comprehension, with the attainment of advanced literacy as a result.

1.1.2 Early reading instruction

In the literature on early reading instruction, it is stated that such instruction should focus on three core elements: phonemic awareness, mastery of the alphabetic principle and the

development of fluent reading (Houtveen, Brokamp, & Smits, 2012; Simmons & Kame'enui, 1998).

Phonemic awareness and the alphabetic principle are both taught as part of kindergarten reading instruction and thus early literacy development. Phonemic awareness requires listeners to hear, identify and manipulate phonemes which are the smallest units of sound which make a difference in meaning. Phonemic awareness can be further divided into phonemic analysis and phonemic synthesis (Wagner, Torgesen, & Rashotte, 1994). Phonemic analysis is the ability to separate spoken words into their respective parts (i.e., phonemes); phonemic synthesis is the ability to merge phonemes together to produce a spoken word. By the end of kindergarten, children should be able to detect, manipulate and analyze the auditory aspects of spoken language. This includes the ability to distinguish and segment words, syllables and phonemes. In addition, children should be able to connect some written or printed letters (i.e., graphemes) with their corresponding phonemes.

Explicit instruction of the alphabetic principle is typically undertaken during the first half of grade 1. Children are asked to name letters and blend them into words, which can be slow and laborious at the beginning of grade 1 but improve as the children gain more experience with the decoding of words and progress towards more rapid word analysis.

During the second half of grade 1, pupils are taught to read a text fluently. Fluent reading is the ability to read a text rapidly, correctly, and using the right intonation. Fluent readers can process punctuation, understand where stress must be placed and know when to pause when reading a text (Houtveen, Brokamp, & Smits, 2012).

1.1.3 Policy regarding the goals of reading instruction

While the Dutch government gives schools considerable leeway with regard to their instructional approaches and curricula, intermediate and core learning goals to be met by the schools have also been laid out (Ministry of Education, Culture and Science, 2006). For each of the core (i.e., central) goals, a specific learning route and intermediate goals are spelled out. The learning route stipulates what all children should be offered during the primary school period. The intermediate learning goals are marked steps for the pupil's development (i.e., milestones to be reached). The intermediate goals are described in such a manner and in sufficient detail that schools can realize all of the required learning routes and central goals during the eight years of education they provide.

The 'beginning reading and writing' learning route is divided into two sub-routes: 'beginning reading and writing, start' and 'beginning reading and writing, continuation.' By the end of the 'beginning reading and writing, start' sub-route (after the first half of grade 1), the following intermediate goals should be met: pupils should know most letters and thus the links between most graphemes and phonemes; pupils should be able to read and write phonologically regular words of the type Consonant Vocal (CV), Vocal Consonant (VC) and Consonant Vocal Consonant (CVC) (Verhoeven & Aarnoutse, 1999). The 'beginning reading and writing, continuation' sub-route stands central in the second half of grade 1. By the end of grade 1, the following intermediate goals should be met: pupils can read and write phonologically regular words of the type CCVC, CVCC and CCVCC; they can read short words with different spelling patterns and also polysyllabic words; they can call upon a wide variety of word-identification techniques and they can automatically recognize many words (Verhoeven & Aarnoutse, 1999).

1.1.4 The instructional programme Learning to Read Safely

Primary school teachers can call upon different programmes to teach reading. A programme that is very often used in the Netherlands is the Veilig Leren Lezen programme (Learning to Read Safely progamme) (Mommers, Verhoeven, Koekebacker, Van der Linden, Stegeman, & Warnaar, 2003).

The Learning to Read Safely programme is organized thematically (e.g., 'my class' and 'my body') and composed of 12 units. During the first 6 units, which are taught during the first half of grade 1, phonemic awareness is stimulated; the pupils learn about the alphabetic principle; and the pupils learn to accurately read simple CVC, CCVC and CVCC words. During the second 6 units, which are taught during the second half of grade 1, the pupils are taught to read sentences and text fluently.

Each unit in the Learning to Read Safely programme has a manual in which 'what' to learn and 'how' to teach it are systematically described for either 20 days (20 lessons for unit 1) or 15 days (15 lessons for units 2 through 12, respectively). Each lesson consists of four phases. During the first Introduction phase, the pupils are motivated to read by, for example, having a story read to them. During the next phase, the Instruction phase, the exercises for the lesson are presented; new learning content is introduced; relevant activities are demonstrated and the teacher checks that the pupils understand the new learning content. In the third phase, the Practice phase, the pupils are given the opportunity to independently apply what has been learned; the teacher provides feedback and correction as needed; and poor readers may be given additional guidance and explanation of the new learning content by the teacher. In the fourth and final Review phase, the teacher and pupils reflect upon the objective(s) of the lesson, the past week or the unit in general.

The Learning to Read Safely programme further differentiates between pupils for purposes of instruction via convergent differentiation. Convergent differentiation means that minimum goals are set for the group as a whole and differences are handled by practicing different activities with different children. The aim of convergent differentiation is for all of the children in the group to remain motivated to learn and, as a result of this, achieve at least the minimum set of goals.

At the start of grade 1, in the Learning to Read Safely programme a distinction is made between readers, average readers, and poor readers. Each group is given different activities adapted to their own reading level, next to the common activities for all groups. Pupils who are able to read a simple book independently, at the beginning of the year, are classified as readers. Pupils who are able to follow the instructions without serious problems are classified as average readers. And pupils who have serious problems with reading are classified as poor readers. During the first two phases of a lesson (i.e., the Introduction and Instruction phases), all groups are jointly involved. During the Practice phase, however, the readers, average readers and poor readers are given different activities. Poor readers practice with the necessary learning content under the guidance of a teacher, average readers practice independently applying what has been learned and readers work with materials specially developed to stimulate not only beginning reading but also reading comprehension. During the final Review phase, the lessons are again completed jointly by the readers, average readers and poor readers.

1.1.5 Different instructional approaches in learning to read

Most children acquire reading skills quite easily. However, reading difficulties and disabilities are an unfortunate reality for some (e.g., McGill-Franzen, Ward, Goatley, & Machado, 2002; Vellutino & Fletcher, 2005). Alternative instructional approaches may thus be called for and more effective when used with specific pupils or groups of pupils, but which instructional approaches are best suited for which pupils or groups of pupils?

This question is related to major questions confronting theories of learning and instruction today, namely: Is knowledge best provided or generated (see Kameenui & Carnine, 1998; Rosenshine, Meister, & Chapman, 1996) and do children learn better individually or together (Bennett, 1987; Slavin, 1990)? Several researchers have tried to overcome the dichotomies implied by these questions. The first reason for these attempts to go beyond these questions lies in the uni-dimensionality of these theoretical dichotomies. Instructional approaches cannot adequately be depicted on one single dimension. Secondly, to describe instructional approaches in real classroom practices in absolute terms of either providing or generating is not realistic. The same holds true for the dichotomy learning individually or learning together. Actually, these kinds of distinctions are theoretical metaphors for teaching and learning. According to Sfard (1998), it seems that the most powerful research is the one that stands on more than one metaphorical leg. Therefore, it seems more fruitful to look at instructional approaches as more complex interaction patterns to be described on multiple dimensions. These patterns originate in theories of teaching and learning but need to be adapted to the specific classroom situation and the curricular content. A third and final reason to go beyond these kinds of uni-dimensional and absolute thinking can be found in the literature about curriculum innovation (Stahl, 1999). In the innovation of reading instruction the pendulum has swung back and forth between stimulating spontaneous learning activities of children' versus 'stimulating learning activities of children with teacher-structured exercises of rules'.

Against this background several researchers turned to the 'Guided Co-Construction'(GCC) of knowledge as a promising alternative (see Hardman 2008; Mercer, 1995; Terwel, van Oers, van Dijk, & van den Eeden, 2009). Such an instructional approach is also sometimes referred to as co-elaboration, co-construction or guided reinvention (Brown & Palincsar, 1989; Dewey, 1943; Freudenthal, 1991). More recently, from a socio-cultural perspective, McCaslin and Burross (2011) also mentioned the importance of guided elaboration in instructional settings that include and go beyond providing knowledge and skills. The empirical study of Hoek, Terwel and van den Eeden (1997), Prinsen, Terwel, Zijlstra and Volman (2013), van Dijk, van Oers and Terwel (2003), van Schaik (2010) and Keijzer (2003) into the effects of 'guided elaboration' can also been seen in this research tradition. These studies showed positive effects of co-operative learning. However, these results are generally derived from the teaching of subjects like: mathematics (van Dijk, 2002; Hoek, 1998), physics (van Schaik, 2010) and world orientation (Prinsen, 2007).

Within the context of this dissertation we were interested in the effects of two instructional approaches, a providing approach: Direct Instruction (DI) and a approach with a cooperative learning component: Guided Co-Construction (GCC). DI or the direct provision of knowledge has been shown to be an effective instructional approach for teaching children to read and particularly children at risk for reading difficulties (Adams, 1990; 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.

1.1.6 Direct Instruction

Direct instruction (DI) has been studied in several domains of teaching, including mathematics and language instruction. DI is a model for teaching which utilizes carefully planned lessons designed around small learning increments. The instruction is highly structured and describes or even scripts classroom activities in considerable detail.

This model was created by Engelmann and colleagues in 1964 at the University of Illinois Institute for Research on Exceptional Children. The first implementation of the model was known as Direct Instruction System for Teaching and Remediation (DISTAR), which consisted of programs that addressed reading, language, and math (Magliaro, Lockee, & Burton, 2005). In 1986, Rosenshine elaborated on DISTAR model and encouraged teachers to present material in small incremental steps, regularly pause to check for understanding and aim to have active participation from all pupils (Ryder, Burton, & Silberg, 2006). The assumptions underlying these DI methodologies are that DI can prevent or eliminate mis-

conceptions during the learning process, enable accelerated learning and foster efficient learning. Each component of a task associated with a particular target skill or behaviour is taught by the teacher who can also model the skill or behaviour, provide opportunities for practice and feedback, and determine when the skill or behaviour needs to be retaught (Ryder, Burton, & Silberg, 2006).

The Direct Instruction approach has been operationalized in many reading programs varying from highly structured, teacher centered approaches to more interactive variants of Direct Instruction (Stahl, 1999). DI as it is conceived of in this study is not a lecture approach but, rather, an instructional model with a focus on the interaction between teachers and pupils.

1.1.7 Guided Co-Construction

According to Vygotsky (1978), the interaction between a learner and a more knowledgeable other, such as a parent or teacher, is one of the greatest influences on the development of the learner and his or her thinking. In what has come to be known as the social constructivist approach to learning, it is further assumed that the classroom discourse is not effective unless pupils play an active part in the discourse and thus in their own learning (Barnes & Todd, 1995; Wells, 1999). In this way, differences between pupils can be called upon to promote this process, and the learning content is not only prescribed ahead of time but also allowed to emerge from the discourse between pupils with other pupils and their teachers.

Closely related to Vygotsky's work is Bruner's (1985) concept of 'scaffolding'. Scaffolding is the provision of support which is tailored to the specific needs of the pupil with the intention of helping the pupil achieve his or her learning goals (Sawyer, 2006). Not surprisingly, scaffolded instruction has been shown to promote greater learning than non-scaffolded instruction. It is therefore frequently among the aims of education but not always easy to realize in light of large numbers of pupils and lack of manpower. And alternatives have thus been sought.

Mercer (1995) was probably the first to speak of the 'guided construction of knowledge.' Later Hardman (2008) was the first to use the term 'guided co-construction' within the context of language education.

For beginning reading, guided co-construction can be understood to entail the following core elements. 'Guided' refers to the explicit role of the teacher for whole-class instruction and the scaffolding of the learning of pupils in either groups or individually. 'Co-' refers to the cooperative aspect of the learning of reading. And 'construction' refers to the development of insight, skill and behaviour on the basis of accumulated knowledge and experience. 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).

1.2 Overarching question

The development of beginning reading, especially word recognition, has been widely studied both nationally and internationally with the following factors identified as important determinants of early word recognition: phonemic awareness, letter knowledge and rapid automatic naming of letters, digits, objects and colours, writing of letters (Aarnoutse, van Leeuwe, Voeten, & Oud, 2001; Bowers & Swanson, 1991; Hansen & Bowey, 1994; Näslund & Schneider, 1996; National Early Literacy Panel, 2008; Verhoeven & Van Leeuwe, 2003; Wagner & Torgesen, 1987). This information, however, does not provide us with an answer to the question of which instructional approach is most suited to stimulate the development of children's word recognition in general or which instructional approach may be best suited for a particular pupil or group of pupils. In other words, the development of children's early word recognition has yet to be considered in connection with the type of instruction which they receive during the early school years, which is what we therefore set out to do.

Within the context of the present research, the overarching question is: How does the word recognition of primary school pupils develop during the first grade with a Direct Instruction approach and with a Guided Co-Construction approach and what factors appear to influence the development of the children's word recognition? This overarching question encompassed the following three research questions:

1. Which instructional approach (DI or GCC) more effectively stimulates the development of word recognition in grade 1?
2. Which kindergarten pre-literacy skills appear to be important for the development of children's word recognition?
3. How does the word recognition of first grade pupils develop?

1.3 Outline of the studies

In Chapter 2 we determined which form of instruction best stimulates the development of children's early word recognition, Direct Instruction or Guided Co-Construction. For this study a field experiment with a pre-test/post-test control group design was undertaken. Two variants of the reading programme Learning to Read Safely were implemented among two different groups of pupils: a variant with a predominantly DI approach (i.e., the

ch1

control group) and a variant with a more GCC approach (i.e., the experimental group). To determine which approach was most effective, the development of 178 children's word recognition throughout the course of grade 1 was compared. Our more specific question was: Is it 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.

In order to substantiate the insights provided by the field experiment into the effectiveness of the different instructional approaches, the outcomes should be seen in the light of the question to what extent the instructional programs were implemented in classroom practices. In Chapter 3 the quality of the implementation of the two instructional approaches (DI and GCC) was therefore assessed. The research question was whether the Direct Instruction and Guided Co-Construction approaches were implemented as intended. That is, did the activities of the pupils and teachers in the DI group show more characteristics of DI than the activities of the pupils and teachers in the GCC group? And conversely, did the activities of the pupils and teachers in the GCC group show more GCC characteristics than the activities of the pupils and teachers in the DI Group? To answer these questions, we carefully described the implementation process for the two approaches (i.e., DI versus GCC), collected observational data from the DI and GCC groups, and compared the quality of implementation for the two groups.

In Chapter 4 we determined whether children's word recognition skills could be predicted. It has been suggested that there are several cognitive prerequisites that contribute to learning to read (Bowers & Swanson, 1991; Hansen & Bowey, 1994; Näslund & Schneider, 1996; National Early Literacy Panel, 2008; Wagner & Torgesen, 1987). One of the major antecedents of reading skills is a child's phonemic awareness. Another important prerequisite is letter knowledge, which has been shown to provide a basis for understanding the alphabetic principle. It has also been suggested that naming speed is an important antecedent of learning to read (Verhagen, 2009). Therefore, the research question of the third study was: Is there an effect of phonemic awareness, letter knowledge and naming speed in kindergarten on children's word recognition skills after six month of reading instruction in grade 1?

In the fifth and final Chapter the development of word recognition of first grade pupils was examined. According to some researches, children with different levels of ability can be expected to show different patterns of reading performance over time (e.g., Stanovich, 1986; Williamson, Appelbaum, & Enpanchin, 1991). Therefore, the research questions of the fourth study were: How does word recognition develop during grade 1? How does word recognition develop during grade 1, for pupils initially classified as poor and non-poor readers, and are pupils tied to one category, i.e. poor reader or non-poor reader, during grade 1?

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.

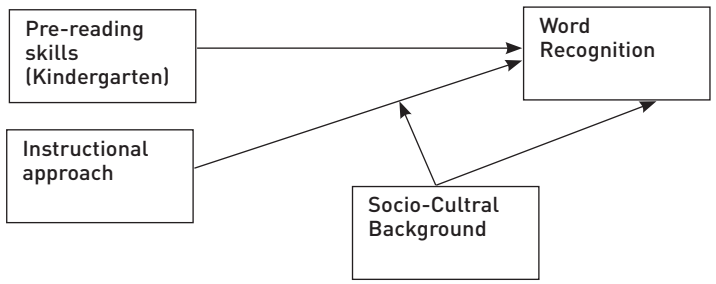
2.2.5 Data analyses

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			

ch2

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

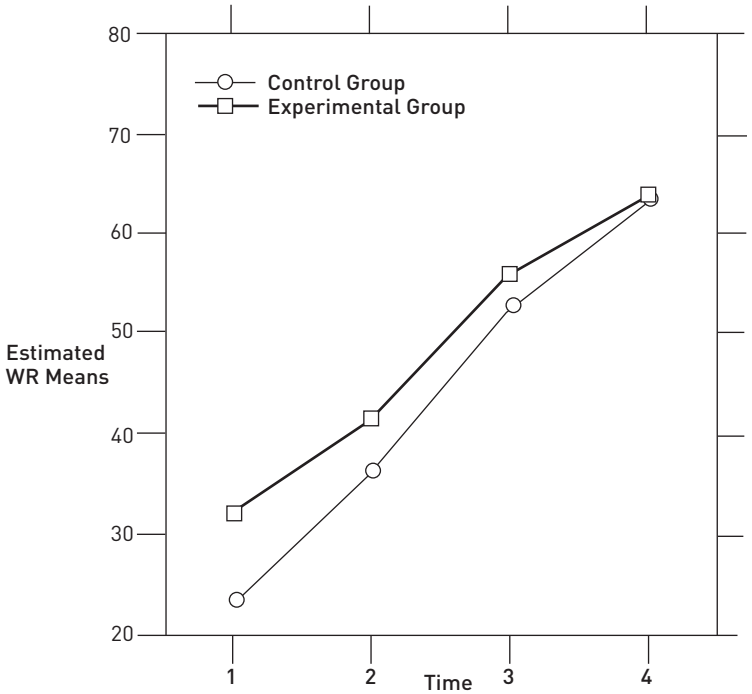


Figure 2: Development of word recognition in control versus experimental groups.

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.

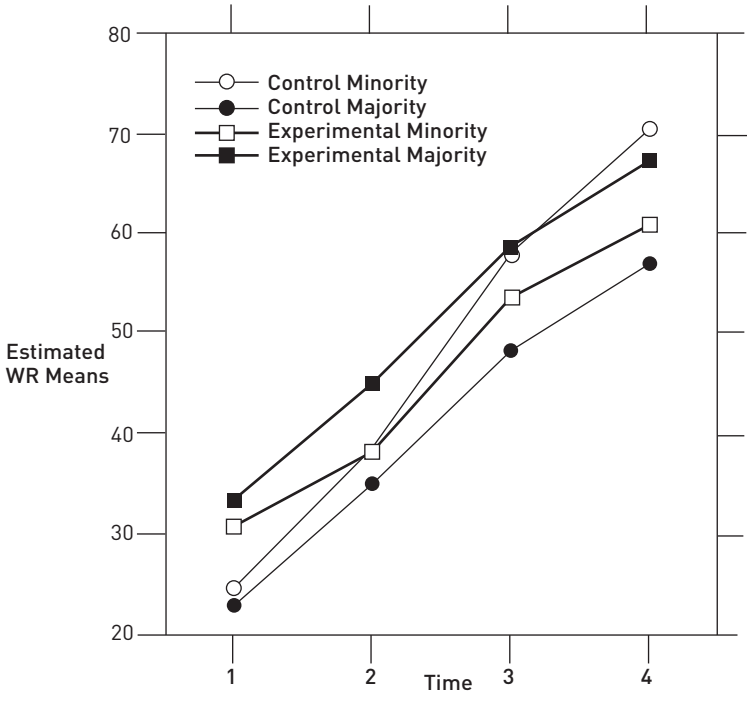


Figure 3: Development of word recognition according to socio-cultural backgrounds of students in control versus experimental groups

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 comprehensive 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.

Chapter 3

The implementation of two instructional approaches: Guided Co-construction and Direct Instruction in a beginning reading programme[2]

[2]Snel, M.J., Terwel, j., Aarnoutse, C.A.J. & van der Veld, W.M. (submitted). The implementation of two instructional approaches: Guided Co-construction and Direct Instruction in a beginning reading programme.

Abstract

In this study, the quality of implementation for a beginning reading programme using two different instructional approaches was assessed. The implementation of the beginning reading programme using Direct Instruction was analyzed and the implementation of the same programme using Guided Co-Construction. Our specific research question was whether the Direct Instruction and Guided Co-Construction approaches were implemented as intended. That is, did the activities of the pupils and teachers in the DI group show more characteristics of DI than the activities of the pupils and teachers in the GCC group? And conversely, did the activities of the pupils and teachers in the GCC group show more GCC characteristics than the activities of the pupils and teachers in the DI Group? In order to assess the quality of implementation, we conducted direct observations and analyzed these both qualitatively and quantitatively. Separate, illustrative examples of DI and GCC could be identified and described in terms of the qualitative characteristics of DI and GCC depicted in the video taped episodes.The sums of the systematic time-sampled observations of teacher and pupil activities in the DI and GCC conditions were next calculated and found to differ as expected. These qualitative and quantitative differences show the implementation of the two approaches to clearly differ in terms of actual classroom practice. Both the teacher and pupil activities reflect the differences which we expected to see for a Direct Instruction versus Guided co-construction approach to beginning reading instruction.

3.1 Introduction

Literacy research is increasingly being called upon by policy makers and school managements to inform their decision-making but also by practitioners to determine how instruction should be approached to maximize pupil outcomes (Pressley, Graham, & Harris, 2006). Within this context, research by Snel, Terwel, Aarnoutse and van Leeuwe (2012) has contributed to the growing body of knowledge examining the effectiveness of two approaches to instruction, namely: Direct Instruction (DI) versus an approach with a cooperative learning component, Guided Co-Construction (GCC).

In their research, Snel et al. (2012) compared two versions of the standard Dutch beginning reading and spelling programme Learning to Read Safely (Mommers, Verhoeven, Koekebacker, van der Linden, Stegeman, & Warnaar, 2003). A predominantly DI approach was adapted to form a GCC approach, and the two instructional versions of the programme were then implemented with separate groups of pupils in first grade: an experimental GCC group and a control DI group. To determine which instructional approach was most effective, Snel et al. assessed the word recognition of the pupils. Word recognition is the process of converting a sequence of written graphemes into sounds for the identification of a word (Coltheart, 1978). Snel et al. found the pupils in the GCC group to outperform the pupils in the DI group when the development of their word recognition was compared across first grade. However, the initially better performance of the GCC group attenuated over time with better performance found for the DI group by the end of first grade. When Snel et al. also examined the roles of socio-cultural background variables in the responding of the pupils to the instructional approaches, they found majority Dutch pupils to benefit most from GCC but minority non-Dutch pupils to benefit most from DI. Furthermore, the minority pupils in the DI group showed the greatest word recognition progress.

Whether or not the results of the study by Snel et al. (2012) can be attributed to the instructional approaches used depends on the instructional approaches being implemented as intended. The research question in the present study was therefore: Have the Direct Instruction and Guided Co-Construction versions of the beginning reading programme Learning to Read Safely been implemented as intended? In order to answer this question, we described the implementation process for the two versions of the programme and compared the activities of the pupils and teachers in the two groups. If DI characteristics appeared more often in the activities of the pupils and teachers in the DI group than in the GCC group and, conversely, more GCC characteristics appeared in the activities of the pupils and teachers in the GCC group than in the DI Group, then it can be concluded that the implementation of the instructional approaches occurred as intended.

3.1.1 Implementation

Implementation refers to the process of putting an intervention into actual practice (Lendrum & Humphrey, 2012). Research in many fields, including the field of education, has shown interventions to be rarely implemented as designed (Lendrum & Humphrey, 2012). In the field of education, the teacher's task is to do what the formal curriculum prescribes and implement the curriculum as intended — regardless of situational differences (Gresham, MacMillan, Boebe-Frankenberge, & Bocian, 2000). However, educational implementation research shows teachers to implement curricula differently even when the curriculum is the same because of differing pedagogical responsibilities and teaching circumstances (i.e., situational differences) (Fullan, 1991; Fullan & Pomfret, 1977; Goodlad, 1994; Knight, 2001; Snyder, Bolin, & Zumwalt, 1992). An instructional curriculum must always be adapted to the actual situation, which means that implementation is always a matter of degree (Hoek, Terwel, van Hout-Wolters, & van den Eeden, 2003; Revicki, 1981). And the degree or fidelity of curriculum implementation is therefore just as important as any other element of programme evaluation (Durlak, 2010, p. 356). Stated differently: Deviation in the implementation of an intervention and thus the use of an educational curriculum can affect achievement outcomes and thereby confound programme evaluation. Implementation as intended is rare, however, while research suggests that local changes are inevitable and surface-level adaptations can influence not only the efficacy but also the sustainability of an educational intervention (Lendrum & Humphrey, 2012).

To determine the extent of treatment fidelity or, in other words, the extent to which an educational curriculum is implemented as prescribed, the key characteristics of the curriculum must first be articulated. In the present study, we therefore translated the

theoretical concepts underlying the DI and GCC approaches to the teaching of beginning reading into key instructional characteristics and behaviours observable on the parts of teachers and pupils. For purposes of the present study, we were interested in mainly the interactions between teachers and pupils, which can be articulated in terms of asking questions, providing answers and guiding/performing specific learning tasks. We did not analyze all of the characteristics of implementation as this was beyond the scope of the present study. We did not analyze, for example, the social aspects of cooperation which might include taking the perspective of the other in order to understand the perspective of the other (Ivey, 1994) or the specific aspects of verbal interaction which are known to promote effective explanation (Webb, 1992).

3.1.2 Direct Instruction

Direct instruction (DI) has been studied for several domains of teaching, including mathematics and language instruction. DI is a teaching model which emphasizes the use of carefully planned lessons designed to achieve small, specific learning increments. The instruction is highly structured and clearly outlines or even scripts classroom activities in considerable detail. The DI model was created by Engelmann and colleagues in 1964 at the University of Illinois Institute for Research on Exceptional Children. Within the context of beginning reading instruction, Slavin, Lake, Chambers, Cheung and Davis (2009, p. 1406) have 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'. DI addresses both 'what' to teach (i.e., the content of a curriculum) and 'how' to teach (i.e., specific teaching techniques).

DI is further characterized by a large amount of activity on the part of the teacher and a strongly directive role for the teacher. Teacher control and feedback are important indicators of directive teacher behaviour (Biemans, Jongmans, de Jon, & Bergen, 1999). The teacher will ask questions like: 'What is this letter?', 'What word is this?' or 'What sound do you hear at the front of the word 'moon'?. The purpose of these questions is to check that the pupil has understood the course material (Creemers, 1991). The responses of the pupils are predominantly right or wrong. And the teacher guides the pupil to the desired result by giving feedback such as 'Well done!' or 'No, that's not right.'

3.1.3 Guided Co-Construction

DI leaves little room for pupil-initiated activity and, in particular, activities which can shape and direct their own learning processes. In a Guided Co-Construction (GCC) approach to instruction, the teacher no longer plays just the role of knowledge provider but also now guides learning processes. Central to the guided co-construction of knowledge is the teacher talking **with** pupils in both a whole-group context but also small-group and individual contexts in order to stimulate and guide their thinking and learning. 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; Hardman, 2008; Wells, 1999).

Within the context of beginning reading instruction, the GCC approach 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 phonemes, graphemes, words and sentences directly but also eliciting and scaffolding contributions from the pupils themselves (i.e., stimulating the construction of knowledge within a meaningful context). In this interactive process, the phonics repertoire of graphemes and words is not only prescribed but also created by having pupils interact and thereby calling upon the differences between pupils to scaffold their knowledge construction. 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). The teacher plays a central role in guiding the learning process, and the activities of the teacher are aimed mostly at stimulating and activating pupils to share what they already know in such a way that they jointly discover new and not already offered knowledge. The teacher generally poses open questions, like 'What other letters do you know?'or 'What word can you create with that letter?'. These questions can be answered in different ways depending on the experiences of the pupils and their already existing knowledge.

3.1.4 Research question

The present study focussed on the quality of implementation for a beginning reading programme using two different instructional appraoches: a Direct Instruction appraoch and a Guided Co-Construction approach. The research question was whether the Direct Instruction and Guided Co-Construction approaches were implemented as intended. That is, did the activities of the pupils and teachers in the DI group show more characteristics of DI than the activities of the pupils and teachers in the GCC group? And conversely, did the activities of the pupils and teachers in the GCC group show more GCC characteristics than the activities of the pupils and teachers in the DI Group?

ch3

3.2 Methods

3.2.1 Participants and Research design

A total of nine grade 1 classes from eight schools located in or near the city of Utrecht (the Netherlands) participated in this study. The experimental GCC group consisted of four classes for a total of 88 pupils. The DI control group consisted of five classes for a total of 90 pupils. An important condition for inclusion in the study was that the class was using the standard Dutch programme for beginning reading and spelling instruction, namely Learning to Read Safely programme (see below). After intake interviews with the teachers of grade 1, the schools were classified such that different types of schools were represented across the experimental and control conditions; schools with predominantly minority pupils were equally distributed across the two conditions, just as schools with predominantly majority pupils from rural areas around the city of Utrecht. All of the schools were Catholic schools. Prior to the start of the study, the parents of the participating pupils were all asked to give their permission for use of the anonymous reading results from their child, observation of their child and filming of their child in the class for purposes of the present research.

In August at the start of grade 1, the average age of the 178 participants was 6 years and 6 months (SD = 5.1 months). The sample consisted of 51% boys and 49% girls; 61% non-minority Dutch and 39% minority with mostly Turkish or Moroccan parents. There were 11 non-minority pupils and 43 minority pupils with two low educated parents; 98 non-minority and 26 minority with one or two high educated parents.

The pupils from the DI and GCC groups did not differ from each other with regard to the background variables of gender, age or socio-cultural background. The crucial variable for insuring initial comparability of the instructional groups was reading skill. This was difficult to assess at the beginning of first grade because most of the children could not read at this time. We therefore decided to administer a test which is known to predict early reading skill rather well, namely the Naming Speed Digit Test (Aarnoutse, Beernink, & Verhagen, 2010; Scarborough, 1998; van den Bos, Zijlstra, & Lutje Spelberg, 2002; Verhagen, Aarnoutse, & van Leeuwe, 2008; Wagner, Torgesen, Rashotte, Hecht, Barker, Burgess, Donahue, & Garon, 1997; Wolf, Bally, & Morris, 1986; Wolf & Bowers, 1999). The test was administered at the end of kindergarten, and the average scores of the children in the DI and GCC groups did not differ significantly from each other (diff. = -1.33, $p>0.05$).

3.2.2 Learning to Read Safely programme

In the Netherlands, the programme most frequently used to teach beginning reading and spelling is the Learning to Read Safely programme. The programme is organized thematically (e.g., 'my class' and 'my body') and composed of 12 units. During the first 6 units, which are covered during the first half of the school year, the three steps involved in 'the fundamental reading operation' are covered: 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) phonemic synthesis or the merging of phonemes together; and 3) the assignment of meaning to the word being read. After unit 6 most of the Dutch graphemes are offered. In units 7 through 12 and thus during the second half of the school year, pupils are taught to read text fluently (i.e., to automatize their reading skill). This entails the linking of not only graphemes to phonemes but also letter clusters to syllables.

The Learning to Read Safely programme has a manual to accompany each unit. The manual describes 'what' should be taught on 20 days for the first unit and 15 days for the other units together with 'how' the material is best taught. Each day's lesson has a predominantly DI character and consists of four lesson phases, which are very similar to the six DI phases distinguished by Rosenshine and Stevens (1986). The four phases in the Learning to Read Safely programme are as follows. First, there is the introduction phase aimed at motivating the pupils to read. This might be reading a story to them and is very similar to the 'review phase' of DI. The introduction phase is followed by the instruction phase. During this phase the exercises for the lesson are presented; new learning content is introduced; relevant activities are demonstrated and the teacher checks that the pupils understand the new learning content. This phase is very similar to the 'presentation phase' of DI. The practice phase is third. During this phase, pupils are given an opportunity to independently apply what they have learned; the teacher provides feedback and correction as needed; and some pupils (like those with poor reading skills) practice with the teaching materials under the guidance of the teacher. This phase is very similar to the 'guided' and 'independent practice' phases of DI. The fourth and final lesson phase in the Learning to Read Safely programme is the review phase. In this phase, the teacher and pupils reflect upon the aims of the current lesson, the aims of the past week or the aims of the monthly unit. This review phase is similar to the weekly and monthly 'review phase' used in DI.

3.2.3 Instruction manuals

For purposes of the present study, two different manuals were used. The DI group used the manual which accompanies the Learning to Read Safely programme and was designed to be executed step-by-step (i.e., constitute a form of DI). An experimental teaching-coaching manual was developed on the basis of the original Learning to Read Safely programme manual and used in the GCC condition in the present study. Those exercises from the original DI manual which allowed translation to GCC were adapted, which meant that not all exercises were adapted from the DI manual for use in the GCC manual. The role of the teacher was often changed from a knowledge provider to a guide for ongoing learning processes. And instead of pupils simply listening to the teacher and/or watching a presentation and answering closed questions which therefore fixed answers, exercises were translated

to have pupils listen to instructions which encourage them to discover and answer questions which could be answered quite differently — depending on the experiences of the pupils and their prior knowledge.

The exercises in both conditions and thus both manuals were designed for use in either a whole class setting or a small group setting but they nevertheless differed from each other in the two conditions. An example of the difference between the DI and GCC approaches to an exercise designed for use with the whole class is illustrated in the following.

In the DI group, the graphemes r-v-i-s-p-aa-e are hung out on a string; the graphemes were covered in the previous lesson(s) and should therefore be known to the pupils. In the introduction phase of the lesson, the teacher reviews the graphemes by having the pupils recite them compliments the pupil when this goes well and corrects the pupils when this does not go well. The teacher then introduces a new grapheme and pronounces the associated phoneme while showing the class the related grapheme card. The card is then attached to the string of other graphemes. Next, the pupils work individually in their workbooks on exercises in which the new grapheme stands central. The learning activities of the pupils in this DI exercise are clearly in the hands of the teacher and the workbooks provided. The teacher determines 'what' activity will be done and 'how' the pupils should go about doing it.

In contrast, in the GCC group, the teacher introduces new material but then gives the pupils an opportunity to exchange prior knowledge of the material and any experiences which they have had with it (i.e. peer collaboration). In such a manner, pupils are prompted to construct knowledge of the material right from the start, to learn from each other and to possibly learn more than just the material which has been presented. The same lesson from the DI condition thus looks very different in the GCC condition. In the GCC classrooms, moreover, all of the graphemes to be learned during the course of the year are hung out; only a blank sheet of paper separates those graphemes which have been taught from those to be taught. Furthermore, old material is reviewed at the beginning of each lesson and new material is introduced just as in the DI classrooms but, in the GCC classrooms, the pupils are also then invited to try to read those graphemes which have yet to be taught. And the pupils in the GCC classrooms are invited to mention words which begin with the grapheme being discussed. The mentioned words are written on the blackboard with the target grapheme written in a contrasting colour and, in such a manner, pupils practice with grapheme/phoneme mappings — not only with the lesson materials but also with the words provided by the pupils themselves. The open character of the GCC approach to this exercise encourages pupils to cooperatively discover new knowledge.

An example of the difference between the DI and GCC approaches to an exercise designed for small group use is summarized below.

In the DI group, the pupils create words by combining the set of grapheme cards provided in a binder. The pupils then read the words out loud to each other and thus practice with the presented material.

In the GCC group, in contrast, the binder has cards for all of the graphemes in it — including those graphemes which have yet to be taught. The pupils can thus create words using new graphemes. And reading these words out loud to each other gives the pupils an opportunity to construct new knowledge together.

3.2.4 Teacher training and monitoring of programme implementation

Given that implementation is putting an intervention or programme into actual practice (Lendrum & Humphrey, 2012) and a teacher's task is to implement a curriculum as intended regardless of situational differences (Gresham, MacMillan, Boebe-Frankenberger, & Bocian, 2000), a clearly written manual or protocol and solid training of the teachers is necessary to ensure accurate implementation of the curriculum (i.e., treatment fidelity). Monitoring for accurate implementation is also necessary.

The first author trained the teachers separately in the two conditions to use the respective instructional approaches. The teachers in the DI condition were trained to introduce new material and then have the pupils practice with the material. In contrast, the teachers in the GCC condition were trained to introduce new material and then give the pupils an opportunity to exchange knowledge about the material and discuss their experiences with the new material.

The training of the teachers took place in their classrooms but without the pupils present. The principles underlying the instructional approach, the reading exercises, the role of the teacher and the roles of the pupils were explained and discussed in great detail. If necessary, teaching activities were also demonstrated. The teachers were not informed that they were part of an experimental study.

The teachers were intensively guided by the first author who visited them every six weeks throughout the course of the investigation in order to answer any questions about the programme, deliver the teacher manuals for the upcoming units, discuss the roles of the teacher and the pupils and also monitor just how well the programme was being implemented. During these meetings, the teachers showed themselves to be very positive about their particular instructional approach, enthusiastic about the exercises and pleased with the results.

Monitoring was undertaken to check the accuracy and success of implementation. Training and a clearly written manual or protocol are not enough to guarantee accurate implementation of a programme or intervention. Adequacy of implementation must also be checked and corroborated using a mix of direct and indirect methods which can vary and are usually

specific to the particular programme/intervention. Methods frequently employed for monitoring purposes are questionnaires, interviews and self-reports of behaviour or experiences but also independent observation (Durlak & DuPre, 2008). In the present study, we used: self-report, video recording and observation.

The purpose of the self-reports was to gain insight into what the teachers did and the perceived results of these actions. We therefore asked the teachers to describe their reading activities and the results of these activities everyday (e.g. the words which the pupils created and read aloud). Unfortunately, keeping such self-reports proved to be too much for the teachers. At most, they described the reading activities they conducted.

Video recording and observation were conducted in the classrooms when a new grapheme and word were introduced. This allowed us to gain insight into the adequacy of implementation for both the DI and GCC approaches. On the one hand, the instruction could be characterized as sufficiently directive when the pupils read graphemes and words which had been selected by the teacher and practiced with materials supplied by the teacher. On the other hand, the instruction could be characterized as having sufficient characteristics of GCC when the pupils were clearly given opportunities to contribute their own knowledge or pre-knowledge and personal experiences but also when the pupils were stimulated to read graphemes and words which had not been taught yet. Video recording was not possible in all classes, because filming during the lessons was experienced as stressful by some teachers. In the end, filming (by the first author) occurred only on one occasion in one DI class and one occasion in one GCC class.

Observation was undertaken in each of the nine classes for an average of 5 lessons between November and January. The observation was done by college pupils who were studying education at the Utrecht University of Applied Sciences and intensively trained on a specific method of observation for purposes of the present study. The specific method of observation is called time sampling. This procedure works with fixed units of time (e.g., 30 seconds) in which the behaviour of three target pupils is observed and registered, the behaviour of the teacher is observed and registered and other information on the setting is observed and registered. When this has been completed for the first target pupil, the next target pupil is observed. And when all three pupils have been observed, the observation cycle starts again with observation of the first pupil, teacher and setting followed by observation of the second and third pupils together with the teacher and setting. An observation cycle thus lasts a total of 90 seconds: The teacher is observed three times during this period (i.e., for three pupils) and each pupil is observed once (i.e., for a period of 30 seconds).

3.2.5 Training of the observers

A total of nine observers observed the activities of the teachers. A maximum of three pupils in a class were observed. The observers were not informed of the instructional approach (i.e., DI or GCC) being used in the class. The observers were intensively trained to register the activities of the teacher, the pupils and the setting using standard codes. The training began with an explanation of the observation system. The observers were then trained to score activities which can occur during reading lessons using video training material. Specific codes were available for DI and GCC activities and the observers practiced with the training materials, compared their coding with each other, discussed the differences, and then coded more training material until sufficient coding reliability was achieved. Thanks to this intensive training, the live observation and coding occurred without significant problems.

3.2.6 The observation system

A fairly simple coding of DI, GCC and other activities on the part of the pupils and the teachers was conducted. As can be seen from Table 1, DI activities coded A for pupils and D for the teacher; GCC activities were coded B for pupils and E for the teacher; and other activities were coded C for pupils and F for the teacher.

Table 1: Observation codes

	Code: DI activities	Code: GCC activities	Code: other activities
Pupil	A	B	C
Teacher	D	E	F

Code A was assigned when the pupil performed any, for example, of the following DI activities.

- The pupil is listening and watching a presentation / instruction and answering closed questions. These questions are based on already presended material and are predetermined.
- The pupil reads offered graphemes, words and texts.
- The pupil names words which begin with offered sounds.

Code B was assigned when the pupil performed any, for example, of the following GCC activities.

- The pupil listens to the teacher who explains exercises and answers open questions. These questions can be answered in different ways depending on experiences of pupils and their knowledge/pre-knowledge.
- The pupil reads not offered graphemes, words and texts.
- The pupil names words which begin with not offered sounds.

Code D was assigned when the teacher performed any of the following DI activities (for example).

- The teacher has the pupils read offered graphemes, words and texts.
- The teacher has the pupils name words which begin with offered sounds.
- The teacher asks closed questions.

Code E was assigned when the teacher performed any of the following GCC activities (for example).

- The teacher has the pupils read graphemes, words and texts which have not been offered previously and may be chosen by the pupils themselves.
- The teacher has the pupils name words which begin with sounds which have not been offered previously.
- The teacher asks open questions.

Codes C and F were assigned for other activities performed by the pupil or teacher, respectively. For example, code C was assigned when the pupil left the classroom to go to the washroom; code F was assigned when the teacher sat and read the manual during the lesson.

3.2.7 Data analyses

To determine whether the beginning reading programme Learning to Read Safely with Direct Instruction and the beginning reading programme Learning to Read Safely with a Guided Co-Construction approach were realized as intended and thus accurately and adequately, the video recordings and live observations were analyzed. The video recordings were viewed by the first author and two fragments were selected as particularly illustrative of the characteristics of DI (fragment 1) and illustrative of the characteristics of GCC (fragment 2).

It was next determined if the characteristics of the GCC approach appeared more often in the observations of the GCC group than in the observations of the DI group and vice versa: the characteristics of the DI approach appeared more often in the observations of the DI group than in the observations of the GCC group. This was done by summing the relevant codes per condition and thus for the behaviour of the pupils and the teacher combined per condition. All DI codes were thus summed for the pupils (code A) and teachers (code D) together per condition. All GCC codes were similarly summed for the pupils (code B) and teachers (code E) together per condition. The totals could then be compared per condition. The observational data from the first lesson were not included in the analyses because it was assumed that the teacher, pupils and observers had to get accustomed to each other and that the behaviour occurring at this time may not be representative and/or coded reliably. Neither the number of observations conducted per classroom nor the length of the observations were the same per classroom. The absolute data were therefore converted into percentages for all subsequent analysis.

The inter-observer agreement was determined by having two observers independently observe and register the behaviour of the same pupils and teacher in the same setting. The agreement between the observations (i.e., codes for the same observation cycle) was then determined and the Cohen's Kappa calculated to indicate the degree of agreement.

3.3 Results

3.3.1 Video fragments

In the following, the video fragments which are illustrative of the characteristics of DI and GCC, respectively, are first described. The results for the observational coding of the characteristics of DI and CGG in the two fragments are then presented.

Fragment 1: Characteristics of DI

Setting:
The pupils sit in small groups together, without interaction between them. The teacher stands in front of the class and holds a word card up with the word 'duif' [pigeon] on it.

Teacher: 'Rian'

Rian: 'Duif' [pigeon]

Teacher: 'and if I put this letter on?'

The teacher places a card with the letter 'k' over the card with the letter 'd' to create the word kuif [mohawk hairstyle].

'Fabienne?'

Fabienne: 'Kuif'.

Teacher: 'Kuif'.

The teacher nods.

Fragment 1 depicts the following DI characteristics. The teacher holds up a word card and calls upon a pupil to read it. Rian must then read the word out loud. This is an unambiguous exercise. Next, the teacher asks a question for which 'kuif' is the only answer. The teacher then gives direct feedback on the correctness of the reading by repeating the word 'kuif' and giving an affirmative nod.

Fragment 2: Characteristics of GCC

Setting:
The pupils sit in co-operative groups of 4 or 5 pupils together. All of the graphemes to be presented during the year are hanging in the classroom. Only a blank sheet of paper hangs between those graphemes which have already been taught and those graphemes which still have to be taught. A pupil stands up and using the pointing stick to refer to a grapheme which she already knows but has yet to be taught to the class: ui.

Teacher: 'Nikita, what do we call this letter? '

Nikita: 'Ui'.

Teacher: 'Do you know a word with that letter? That's really smart.

Nikita: 'Buik'. [belly]

Teacher: 'Right. In which position is the letter?'

Nikita: 'In the middle'.

Teacher: The teacher writes the word 'buik ' on the blackboard.

Teacher: 'Okay, you may sit down again'.
'You may consult in your groups. '

The pupil with the pointing stick sits down and the pupils then discuss words which they know with the letter'ui' in their groups.

Teacher: 'Okay, now stop with the consultation. Shhhh...Which word do you know, Jord? '

Jord: 'Huilen'. [cry]

Teacher: 'Huilen'. And what is the position of the letter?'

The teacher writes the word 'huilen' on the blackboard.

Fragment 2 depicts the following GCC characteristics. Nikita is asked to point to a grapheme which has not been taught yet but she happens to already knows. This is obviously an open question because it can be answered in different ways depending on the child' prior knowledge or pre-knowledge. The pupils are then given an opportunity in their small groups to name words which they know with the grapheme 'ui'. The threshold for the

exchange of knowledge is low in the small groups and the pupils exchange what they know and learn from not only each other but also more than what has been taught to date: They thus construct what is new knowledge for at least some of the pupils together. Jord mentions the word 'huilen', which has not been previously mentioned in the class, for example.

3.3.2 Comparison of GCC and DI characteristics in two instructional conditions

The frequencies of the DI and GCC characteristics registered for the teachers and pupils were next examined. Table 2 shows the teacher registrations in the DI and GCC groups. Table 3 shows the pupil registrations in the two groups. In both tables, the distributions of the registrations differed significantly across the conditions, chi-square (1, N = 178) = 241, p = .00.

Table 2: Teacher registrations in the DI and GCC groups, respectively.

Registrations	Condition		Total (%)
	DI (%)	GCC (%)	
DI registrations	2001(73.3)	729 (26.7)	2730 (100)
GCC registrations	505 (38.3)	813 (61.7)	1318 (100)
Total registrations	2753 (68)	1295 (32)	4048 (100)

Table 3: Pupil registrations in the DI and GCC groups, respectively .

Registrations	Condition		Total (%)
	DI (%)	GCC (%)	
DI registrations	2041 (75.3)	670 (24.7)	2711 (100)
GCC registrations	550 (43.1)	725 (56.9)	1275 (100)
Total registrations	2762 (69.3)	1224 (307)	3986 (100)

Tables 2 and 3 show DI activities to predominate in the DI condition for both the teacher (2001 or 73.3%) and pupils (2041 or 75.3%); only a minimum of DI activities are registered in the GCC condition for both the teacher (729 or 26.7%) and the pupils (670 or 24.7%). When the registration of GCC activities is considered, GCC activities are prevalent in the GCC

condition for both the teacher (813 or 61.7%) and the pupils (725 or 56.9%) and to lesser degree in the DI Group for both the teacher (505 or 38.3%) and the pupils (550 or 43.1%).

The inter-observer agreement was determined by having two observers independently observe and register the behaviour of the same pupils and teacher in the same setting. The Cohen's Kappa for the coding of teacher activity (159 observations) and for the coding of pupil activity (159 observations) were calculated. They were found to be 0.7 and 0.7, respectively, which shows the behaviour of the teachers and pupils to be registered as 'adequate to good'.

3.4 Conclusions and Discussion

In this study, the quality of implementation for a beginning reading programme using two different instructional approaches was assessed. The implementation of the beginning reading programme using Direct Instruction was analyzed and the implementation of the same programme using Guided Co-Construction. Our specific research question was whether the Direct Instruction and Guided Co-Construction versions of the beginning reading programme Learning to Read Safely were realized as intended and thus adequately.

3.4.1 Conclusions

In order to assess the quality of implementation, we gathered direct observations and analyzed these both qualitatively and quantitatively. Separate, illustrative examples of DI and GCC could be identified and described in terms of the qualitative characteristics of DI and GCC depicted in the video taped segments.The sums of the systematic time-sampled observations of teacher and pupil activities in the DI and GCC conditions were next calculated and found to differ as expected. These qualitative and quantitative differences show the implementation of the two versions of the beginning reading programme to clearly differ in terms of actual classroom practice. Both the teacher and pupil activities reflect the differences which we expected to see for a Direct Instruction versus Guided co-construction approach to beginning reading instruction. There was, of course, some overlap between the two approaches. Implementation is always a matter of degree, and certain DI and GCC activities are essential and thus bound to occur in virtually any instructional approach.

3.4.2 Discussion

Although the results of this study confirm the expected (and intended) differences in the classroom practices observed in the DI and GCC conditions, we cannot yet conclude that a

causal relation exists between the implementation of a particular instructional approach and a specific learning outcome as documented by Snel, Terwel, Aarnoutse and van Leeuwe (2012). The results of the present study and the theoretical framework on which it draws, nevertheless give us reason to believe that certain learning outcomes can be attributed to classroom practices and, more specifically, the approach adopted for teaching beginning reading.

At this point, some possible room for improvement in the degree of implementation of the instructional approaches should be mentioned. While we worked to ensure implementation fidelity by using carefully developed manuals and giving the teachers a solid training on the particular instructional approach, it is still possible that this training could be improved with the involvement of the teachers in the design of the curriculum, the development of the instructional approach and the design of the exercises. Not only would the teachers have a better understanding of the approach and the accompanying materials in such a manner; they might also have a stronger commitment to using and thus implementing the instructional approach. In addition, preparatory practice on the use of the instructional approach with pupils might also improve the implementation of the instructional approaches used in the present experiment. The number of video recordings and number of live observations in the class were limited. It was difficult to get permission to film in the classes; many of the teachers did not want to be filmed and only those teachers who were more acquainted with the principlan investigator (i.e., the first author) gave permission in the end. It was easier to get permission to observe in the classrooms than to film, but observation during the lessons was experienced as stressful by some teachers. For this reason, we only observed in the classroom during the first half of the school year.

To compensate for the lack of observation during the second half of the year, we conducted regular interviews with the teachers at this time. The interview results showed the teachers and pupils to perform the exercises as intended in each of the conditions in the experiment and also that the teachers were satisfied with the reading results in each condition.

In sum, both qualitative and quantitative analyses of the observational data collected for the present study show the DI and GCC approaches to the teaching of beginning reading to be implemented as intended. Significant differences in the activities observed for both the teachers and the pupils occurred in the expected direction. DI activities occurred more often in the DI group than in the GCC group and GCC activities occurred more often in the GCC group than in the DI Group.

The results of this study have some immediate theoretical and practical implications. The theoretical contrast of 'providing' versus 'generating' can now be transcended and referred to as a more realistic contrast between Direct Instruction and Guided Co-Construction. And the two instructional approaches studied here have both been shown to be feasible for teaching beginning reading in real classroom settings (Snel et al., 2012).

Chapter 4

Prediction of word recognition in the first half of grade[3]

[3]Snel, M.J., Aarnoutse, C.A.J. , Terwel, J., van Leeuwe, J.F.J., & van der Veld, W.M. (accepted). Prediction of word recognition in the first half of grade 1. European Early Childhood Education Research Journal.

Abstract

Early detection of reading problems is important to prevent an enduring lag in reading skills. We studied the relationship between speed of word recognition (after six months of grade 1 education) and four kindergarten pre-literacy skills: letter knowledge, phonemic awareness and naming speed for both digits and letters. Our sample consisted of 178 pupils divided over nine classes. In agreement with the literature, we found that all four kindergarten tests were related to speed of word recognition in grade 1. We also performed a multiple regression analysis with a set of background variables and the four kindergarten tests. The model explained 53% of the variance in speed of word recognition. However, only letter knowledge and naming speed for digits had a significant direct effect. Our conclusion is, nevertheless, that all four kindergarten tests should be used to identify children at risk for reading problems.

4.1. Introduction

The importance of reading in our time is evident. Someone who does not master this skill cannot function properly in our modern society. One of the main tasks of primary education is therefore to teach all children to read and write or, in other words, to become literate.

Literacy can be defined as the ability to communicate well in writing and thereby read and write effectively and efficiently. Word recognition is part of literacy and can be defined as the process of converting a sequence of letters into sounds for the identification of a word. This decoding process can initially be quite slow. As words are encountered more frequently, however, word recognition becomes both faster and more automatic. Specific patterns of letters, morphemes and words are directly linked to their representations in the mental lexicon and therefore recognized with considerable speed (Coltheart, 1978).

Extensive research has been conducted on the development of children's word recognition skills and reading instruction (Aarnoutse, van Leeuwe, Voeten, & Oud, 2001; Henneman, Kleijnen, & Smits, 2004; Verhagen, 2009; Verhoeven & van Leeuwe, 2003; Wentink & Verhoeven, 2001, 2004). These studies showed that reading instruction should begin early (i.e., in kindergarten) with preparatory word recognition activities. Factors which are known to influence the development of children's word recognition should be targeted as part of these early activities, e.g. phonemic awareness, letter knowledge and naming speed (Aarnoutse, 2004; Aarnoutse, van Leeuwe, & Verhoeven, 2000, 2005; Beernink, 2002; Verhagen, Aarnoutse, & van Leeuwe, 2006, 2008).

4.1.1 Predictors of word recognition

Several studies have shown a small number of kindergarten pre-literacy skills to be related to children's later word recognition (Bowers & Swanson, 1991; Hansen & Bowey, 1994; Näslund & Schneider, 1996; National Early Literacy Panel, 2008; Wagner & Torgesen, 1987). Among the kindergarten skills are: phonemic awareness or the ability to detect, manipulate and identify phonemes; letter knowledge or familiarity with the names and sounds associated with printed letters; rapid automatic naming of letters and digits or the ability to rapidly name a random sequence of otherwise well-known letters or digits; rapid automatic naming of objects and colours or the ability to rapidly name random sets of pictures of well-known objects or colours; writing of letters and name or the ability to write letters in isolation and write one's own name; and phonological memory or the ability to remember spoken information across a short period of time.

In the Netherlands, the most frequently studied predictors of children's early word recognition are: phonemic awareness, naming speed and letter knowledge (e.g. Aarnoutse, 2004; Aarnoutse, van Leeuwe, & Verhoeven, 2000, 2005; Beernink, 2002; Verhagen, Aarnoutse, & van Leeuwe, 2006, 2008).

Phonemic awareness is the ability to hear, identify and manipulate phonemes or the smallest units of sounds which differentiate meaning. Separating the spoken word cat into the three distinct phonemes /c/, /a/, /t/ requires phonemic awareness, which can be further divided into phonemic analysis and phonemic synthesis (Wagner, Torgesen, & Rashotte, 1994). Phonemic analysis is the ability to separate spoken words into their respective phonemes; phonemic synthesis is the ability to merge a number of phonemes together to produce the spoken word.

According to Aarnoutse and Kapinga (2007), simple phonemic awareness skills develop in Dutch — the language spoken in the Netherlands — before the start of formal reading instruction in grade 1. For example, the analysis of the first or last phoneme of a spoken word and the synthesis of Dutch consonant, vowel, consonant (CVC) words is quite well-developed in kindergarten.

Naming speed is the speed with which children name a continuous series of highly familiar items as rapidly as possible. The stimuli are typically letters, digits, colours or pictures of familiar objects, and it is assumed that the naming responses are highly automatized and overlearned (Wolf, Bally, & Morris, 1986). In one of the first studies of the association between serial naming speed and reading ability, Spring and Capps (1974) compared the naming speeds for digits, colours and the pictures of relatively common objects for dyslexic versus non-dyslexic children. The dyslexic children named the items more slowly than the non-dyslexic children. Many studies have since demonstrated the robustness of the association between serial naming speed and word recognition, even after verbal and nonverbal IQ, prior word reading ability, short term memory, articulation rate, speed of processing, letter knowledge and phonological awareness were controlled for (Scarborough, 1998; van den Bos, Zijlstra, & lutje Spelberg, 2002; Wagne, Torgesen, Rashotte, Hecht, Barker, Burgess, Donahue, & Garon, 1997; Wolf, Bally, & Morris, 1986; Wolf & Bowers, 1999).

Naming speed further correlates highly with both single-word recognition (e.g., Spring & Davis 1988; Vellutino, Scanlon, Sipay, Small, Pratt, Chen, & Denckla, 1996; Wolf et al. 1986) and textual reading speed (Bowers, 1993). According to van den Bos et al. (2002), however, naming speed for letters and digits (i.e., alphanumeric naming speed) predicts word recognition better than naming speed for colours and pictures. Moreover, van den Bos, and lutje Spelberg (2010) found naming speed for digits to be a better predictor of word recognition speed than naming speed for letters. They explain this finding in terms of the ambiguity of alphabetic stimuli and the smaller size of the set of digits (i.e., 1 to10) than the set of letters in the alphabet (i.e., a to z).

Exactly how naming speed influences word recognition is still not understood. Cutting and Bridge Denckla (2001) outline three, somewhat conflicting, hypotheses regarding the relationship between naming speed and word recognition. First, naming speed can be assumed to be a component of phonological processing (Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993; Wagner, Torgesen, & Rashotte, 1994). Second, naming speed can be assumed to be fundamental for the development and, in particular, the start of orthographic knowledge (Bowers, 1997; Manis, Doi, & Bhada, 2000; Sunseth & Bowers, 1997; Wolf & Bowers, 1999). Third, naming speed can be assumed to be fundamental to memory span although this hypothesis is less prominent than the other two (see Bowers, Golden, Kennedy, & Young, 1994, for a review).

Letter knowledge at its most basic level is the ability to represent letters which only differ from others in a few distinct ways (e.g., the ability to distinguish d from b or d from p). At a more developed level, letter knowledge is familiarity with the connections between written or printed letters (i.e., graphemes) and their corresponding phonemes (Bowey, 2005). Many studies have shown letter knowledge in kindergarten to be one of the best predictors of children's later word reading ability (e.g., Bond & Dijkstra, 1997; Bowey, 2005; Ehri & Sweet, 1991; Ehri & Wilce, 1987; de Jong & van der Leij, 1999; Lonigan, Burgess, & Anthony, 2000; Scarborough, 1998; Share, Jorm, MacLean, & Matthews, 1984; Wagner, Torgesen, & Rashotte, 1994). Moreover, children who later show reading problems have been found to have **less** knowledge of letters in kindergarten than children without later reading problems (de Jong & van der Leij, 2003).

According to Adams (1990) and Bowey (2005), letter knowledge does not necessarily imply that the relevant connections can actually be used for the recognition of words and thus contribute to later reading skill. Children may master phoneme-grapheme connections but still not understand that these connections must be used to recognize words. Adams (1990) suggests that this is the reason why the learning of isolated graphemes (i.e., letters) in kindergarten does not influence later word recognition while the learning of isolated phonological skills does. The teaching of letter knowledge in combination with phonological skills has nevertheless been found to more strongly affect children's later word recognition than the teaching of just phonological skills, presumably because such learning closely resembles the actual process of decoding and recoding words (Blachman, 2000; Bus & van IJzendoorn, 1999).

4.1.2 Research question

The present study focussed on the prediction of children's word recognition at the end of the first half of grade 1. Based on the reviewed literature we have formulated the following research question: Is there an effect of phonemic awareness, letter knowledge and naming speed in kindergarten on children's word recognition after six month of reading instruction in grade 1?

4.2. Method

4.2.1 Participants

Eight primary schools and a total of nine classes or 178 Dutch pupils who were all learning to read Dutch participated in this longitudinal study. Prior to the start of the study, the parents of the participating pupils gave their informed consent for the use of the anonymous reading results from their child for purposes of the present research. The average age of the pupils at the time of initial testing (i.e., the end of kindergarten) was 6 years and 4 months (SD = 5.1 months). Further, 91 were male (51%) and 87 were female (49%); 109 had a non-minority background (61%) and 69 had a minority background (39%). There were 11 non-minority pupils and 43 minority pupils who had one or two parents with a lower education; 98 non-minority pupils and 26 minority pupils had one or two parents with a higher education.

4.2.2 Measures

Letter Knowledge, Phonemic Synthesis, Phonemic Analysis, Naming Speed Digits and Naming Speed Letters were administered at the end of kindergarten (July). The Word Recognition test was administered in January of grade 1.

All of the measures used in this study were administered by interns in the schools. The interns received special training during several sessions to administer the tests. During these sessions each test, its manual, and its administration were thoroughly discussed and practiced. The tests were administered with individual pupils in a separate and quiet room.

Letter Knowledge. A test developed by Aarnoutse, Beernink and Verhagen (2010) was used to measure passive letter knowledge. The test consists of 23 lists of 23 letters each with the letters **x, y** and **q** excluded and the two letters **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.

Phonemic Synthesis. A test developed to measure a child's ability to reconstruct a word from its constituent phonemes was used to measure phoneme synthesis (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.

Phonemic Analysis. A test developed to measure a child's ability to analyze a pseudo-word into its constituent phonemes was used to measure phonemic analysis (Verhagen & Aarnoutse, 2001). The child is asked to listen to 40 pseudowords and name the first phoneme in words like buin and krontebel on 20 occasions and the last phoneme in words like koes and draap on 20 occasions. A Cronbach's α of .94 was reported by Verhagen and Aarnoutse.

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 and accurately as possible. The child's score is the time required in seconds to name the total of 40 items. Naming Speed for Letters uses the letters o, s, m, p and k because these letters are well known by 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.

Word Recognition. The ability of the child to decode printed words was measured using a task developed by Aarnoutse and Kapinga (2007). The child is presented a card with a list of 100 unrelated words of increasing difficulty and thus ranging from CVC words to multi-syllabic words. The child reads the words aloud as quickly as possible. The test score is the number of words read correctly in 90 seconds. The Word Recognition test was administered after 19 weeks of reading instruction (January). The test-retest correlations mentioned in the manual exceed .86.

4.2.3. Data analyses

In order to study which kindergarten tests significantly predict later word recognition, we conducted hierarchical regression analyses. OLS regression assumes, among other things, that the residuals are independent of each other. This assumption is possibly violated for our sample, which drew upon children from the same class within a school. The standard errors may thus be underestimated leading to an overstatement of significant effects. In order to counter this problem we adopted a fixed effects model with class dummies as fixed effects. A drawback of using a fixed effects model is that inferences cannot be made beyond the groups in the sample. This is not an issue for the present study, because we do not want to generalize beyond grade 1.

Another assumption of OLS regression is that the relationships are linear. Our study involves mostly count variables, which makes linear relationships between variables unlikely. As a consequence of such violation, the estimates and standard errors can be biased. In Table 1, the univariate statistics are presented. It is clear that both skewness and kurtosis are significant for almost all of the variables, hence it is unlikely that our relationships are linear. We therefore decided to create dummy variables for some of independent measures and thereby remedy the problem of possible non-linear relations between the variables. The distribution of phonemic analysis is skewed very much to the right; the distribution of letter knowledge shows two peaks. And the distribution of naming speed letters is extremely skewed to the left. For each of these three variables, a new dummy variable was therefore created. As the cut-off point for each variable, we chose to use the mean. And the dummy variables were coded in such a manner that the 0 category meant poor phonemic analysis, poor letter knowledge, and good (i.e., quick) naming speed for letters.

Table 1: Univariate statistics for predictor and outcome variables used in regression analysis.

	Mean	Median	Variance	Skewness	Kurtosis	Min	Max
Phonemic Synthesis	12.07*	12.50	29.71	-.59*	-.37*	0	20
Phonemic Analysis	32.43*	35.00	69.70	-1.60*	2.38*	0	40
Letter Knowledge	12.24*	13.00	38.51	-.21*	-1.31*	0	21
Naming Speed Digits	44.76*	41.80	194.81	.87*	.75*	18.18	96.78
Naming Speed Letters	59.21*	48.40	1.907.38	5.61*	46.89*	18.00	473.72
Word Recognition	39.52*	38.00	383.64	.80*	.24*	9	99

*** p<.05**

Finally, we tested the hypothesis implicitly put forth by Adams (1990), namely that not only kindergarten letter knowledge is important for grade 1 word recognition but also the combination of kindergarten letter knowledge and phonemic awareness which significantly predicts later word recognition. This was done by adding the following interaction terms to the regression model: letter knowledge * phonemic synthesis; letter knowledge * phonemic analysis; and letter knowledge * phonemic synthesis * phonemic analysis.

4.3.Results

In Table 2, the results of the simple regression analysis are summarized. The three background variables: gender, age and ethnicity did not have an effect on word recognition. The

other predictor variables all had significant effects on word recognition with a considerable amount of the variance in word recognition accounted for by naming speed digits and naming speed letters in particular.

Table 2: Simple regression results with word recognition as the dependent variable.

	Intercept	Effect	R	R2
Gender	38.48*	2.02*	0.05	0.03
Age	27.14*	0.16*	0.04	0.02
Ethnicity	39.19*	0.54*	0.01	0.00
Phonemic Synthesis	20.99*	1.54*	0.43	0.13
Phonemic Analysis$_a$	28.64*	16.55*	0.42	0.12
Letter Knowledge$_a$	30.35*	17.37*	0.44	0.20
Naming Speed Digits	77.86	-0.86*	0.60	0.33
Naming Speed Letters$_a$	45.65*	-19.49*	0.43	0.22

* p<.05

$_a$ These variables are not the original variables, but the dummy variables we created.

The hierarchical regression analysis was conducted next with the addition of dummy class variables and the interaction terms for the predictor variables of letter knowledge, phonemic synthesis and phonemic analysis (see Table 3).

The R2 changes show phonemic synthesis, phonemic analysis, letter knowledge and the naming speed tests to all make significant contributions to the prediction of word recognition.

In the final model, all predictors explain 53% of the variance in word recognition. However, not all variables showed a significant direct effect. In model 6 after the introduction of naming speed digits and naming speed letters, the effects of letter knowledge and naming speed digits remain significant.

Table 3: Hierarchical regression estimates with word recognition as outcome variable.

Model	1	2	3	4	5	6
(Constant)	42.44*	37.45*	24.64*	30.01*	33.39*	76.56*
Class 1	5.40*	5.32*	4.53*	4.99*	5.32*	7.99*
Class 2	-2.80*	-1.18*	1.32*	1.29*	.41*	3.49*
Class 3	-11.88*	-11.72*	-9.32*	-11.36*	-11.02*	-7.77*
Class 4	-3.15*	-1.44*	2.08*	3.63*	2.75*	3.51*
Class 5	-.68*	.99*	-.43*	-.17*	-.62*	-.28*
Class 6	-1.28*	1.42*	1.77*	3.79*	3.51*	3.38*
Class 7	-10.28*	-10.12*	-7.32*	-6.84*	-7.46*	-1.53*
Class 8	-1.84*	-2.06*	-.46*	1.09*	.81*	4.61*
Boy		2.70*	3.21*	4.40*	3.76*	3.96*
Age		.02*	-.08*	-.16*	-.17*	-.31*
Dutch		1.86*	2.4*	2.83*	2.32*	1.15*
Phonemic synthesis			.82*	.54*	.28*	.26*
Phonemic analysis			13.35*	8.33*	8.90*	3.94*
Letter knowledge				12.69*	11.64*	7.04*
Letter knowledge* Phonemic synthesis					.86*	.23*
Letter knowledge* Phonemic analysis					4.20*	7.33*
Letter knowledge* Phonemic synthesis* Phonemic analysis					1.81*	.72*
Naming speed digits						-.59*
Naming speed letters						-4.48*
R2	.08*	.08*	.29*	.36	.38*	.53*
R2 change		0.05*	.21*	.07*	.02*	.16*

* p<.05

ch4

4.4 Conclusions and Discussion

4.4.1 Conclusions

The main purpose of this study was to examine the predictive value of kindergarten tests for word recognition after the first half of grade 1. In line with suggestions from the literature (e.g. Verhagen, Aarnoutse, & van Leeuwe, 2006, 2008) we analyzed the effects of the following kindergarten tests: phonemic awareness, letter knowledge and naming speed We found that all kindergarten tests to be significant predictors of word recognition in our separate regression analyses. The results of the subsequent hierarchical regression analysis, that included all predictors, showed that letter knowledge and naming speed digits are statistically significant predictors of word recognition. The other predictors: phonemic awareness, naming speed letters, letter knowledge * phonemic synthesis, letter knowledge * phonemic analysis, and letter knowledge * phonemic synthesis * phonemic analysis were not statistically significant.

4.4.2 Discussion

In a separate regression analysis used in our study we found that kindergarten phonemic awareness, letter knowledge and naming speed are significant predictors of word recognition. These results correspond with several Dutch studies of children's early word recognition (e.g. Aarnoutse, van Leeuwe, & Verhoeven, 2000, 2005; Verhagen, Aarnoutse, & van Leeuwe, 2006, 2008). However, a hierarchical regression analysis, in which we also added the interaction terms showed that only letter knowledge and naming speed for digits effect children's later word recognition. The other predictors were found not to be statistically significant. Apparently, phonemic awareness and naming speed letters are related to word recognition, which we found in the separate regression analyses. However, these kindergarten tests have no direct effect on word recognition. In addition, we could find no evidence for what Adams (1990) claims, namely that letter knowledge alone (i.e., the isolated learning of graphemes in kindergarten) will not predict children's later word recognition, while the combination of letter knowledge with phonological skills (letter knowledge * phonemic synthesis; letter knowledge * phonemic analysis; and letter knowledge * phonemic synthesis * phonemic analysis) does. The results of our study, however, showed that letter knowledge alone (without the combination with phonological skills) is an important predictor of word recognition.

According to Bond and Dijkstra (1997), Bowey (2005), Ehri and Sweet (1991), Ehri and Wilce (1987), and de Jong and van der Leij (1999), letter knowledge influences word recognition because letter knowledge reflects recognition of the connections between graphemes and their corresponding phonemes (i.e., the link between written and spoken language). Recognition of the grapheme-phoneme link is an important step in the process of converting a sequence of letters into a series of sounds for the identification (i.e., reading) of a word.

The explanation for why naming speed is found to be such a strong predictor of word recognition presumably lies in the common factor: speed. In both the test used to assess naming speed letters and the test used to assess naming speed digits, the pupils are asked to respond as quickly as possible. However, after introduction of letter knowledge into our regression model, the effect of naming speed for letters became nonsignificant. We suspect that this is due to the overlap between the letter knowledge test and naming speed letters test: In both tests, the children must call upon their letter knowledge to name the presented letters.

On the basis of the present findings, we strongly recommend the measurement of children's letter knowledge and naming speed digits at the end of the kindergarten period. Such information allows us to detect stagnation in the development of children's early word recognition skills and therefore initiate early interventions to prevent further stagnation. Exactly how remedial instruction for pupils who fall behind on letter recognition and/or naming speed can improve their skill and facilitate their later word recognition should be examined in future research.

Chapter 5

Convergent or divergent development of word recognition in beginning reading?[4]

[4]Snel, M.J., Aarnoutse, C.A.J., Terwel, J., & van der Veld, W.M. (submitted). Convergent or divergent development of word recognition in beginning reading?

Abstract

In this study we analyzed the development of beginning reading. We have measured word recognition of 178 children 4 times during grade 1. Our research questions were: How does word recognition in general develop during grade 1? How does word recognition develop during grade 1, for pupils initially classified as poor and non-poor readers? And are pupils tied to one category, i.e. poor reader or non-poor reader, during grade 1?

The results of this study showed that the development of word recognition in grade 1 is roughly linear and that poor and non-poor readers show a parallel development of word recognition. In addition, pupils are not tied to the classification made at the beginning of grade 1. About 40% of the poor readers in November manage to become non-poor readers in May, and about 10% of the non-poor readers in November become poor readers in May. Therefore, it is very important that in grade 1 word recognition skills are measured frequently in order to adapt the amount and type of reading instruction to the pupils reading skills.

5.1 Introduction

Teaching children to read is a complex task. The purpose of beginning reading instruction is to help children master the many challenges of the written word, including knowledge of the alphabetic system, the ability to decode new words, a vocabulary that allows words to be read at sight, and the ability to construct, integrate and remember the meanings of words in text. The purpose of this study was to determine the development of first grade pupils with regard to an important aspect of beginning reading, namely word recognition. Word recognition is the process of converting a sequence of letters into sounds for the identification of a word. This decoding process can initially be quite slow. As words are encountered more and more frequently, word recognition becomes more automatic and thereby faster. Specific patterns of letters, morphemes, and words get directly linked to their representations in the mental lexicon (Coltheart, 1978). As a result, the initially arduous process of word recognition gives way to smooth reading and comprehension.

5.1.1. Beginning reading instruction

Beginning reading instruction in the Netherlands starts in grade 1, around the age of six years. In the first half of grade 1 all of the regular grapheme-phoneme correspondences are taught. The instructions are focused on the reading of monosyllabic words of the type Consonant Vowel (CV), Vowel Consonant (VC) and Consonant Vowel Consonant (CVC) (Verhoeven & Aarnoutse, 1999). The second half of grade 1 is focused on the automatization of word recognition skills. During this period, the pupils are taught via repeated practice to recognise words more directly in memory and as a result reading becomes faster and more fluent (Gentry, 2006). At the end of grade 1, pupils can read phonologically regular words of the type CCVC, CVCC and CCVCC; they can read short words with different spelling patterns and also polysyllabic words; they can call upon a wide variety of word-identification techniques and they can automatically recognize many words (Verhoeven & Aarnoutse, 1999).

5.1.2 Beginning reading programme Learning to read safely

Primary school teachers can call upon different programmes to teach reading. The most frequently used programme for beginning reading instruction in the Netherlands is Learning to read Safely (Mommers, Verhoeven, Koekebacker, van der Linden, Stegeman, & Warnaar, 2003). This programme is largely based on phonics, but emphasizes the entire language experience. During the first half of grade 1, the focus of this programme is the three steps in 'the fundamental reading operation': 1) linking graphemes to phonemes in the direction of reading and thus from left to right, which requires the visual analysis of graphemes, linking them to phonemes, and remembering phonemes in sequence; 2) auditory synthesis or the merging of phonemes; and 3) assignment of meaning. After this, during the second half of grade 1, more complex and varied texts are offered and thereby increase the pupils reading fluently.

In the Learning to Read Safely programme a distinction is made between readers, average readers, and poor readers, based on the pupil's reading skills. Each group is given different activities adapted to their own reading level, next to the common activities for all groups. Pupils who are able to read a simple book at the beginning of the year are classified as readers. They are given extra difficult activities that stimulate reading comprehension. Pupils who are able to follow the instructions without serious problems are classified as average readers. They are only participating in the common reading activities. Pupils who have serious problems with reading are classified as poor readers. They are receiving additional guidance of the teacher. The aim of the differentiation is to adapt the instructions to the reading skills in order to keep all pupils motivated and to minimize the number of pupils who fall behind.

5.1.3 Differences in reading development

The performance of poor versus non-poor readers has been studied for decades (e.g. Aarnoutse, Mommers, Smits, & Van Leeuwe, 1986; Jacobson, 1999; Leppänen, Niemi, Aunola, & Nurmi, 2004). The development of poor versus non-poor readers can theoretically be divergent, convergent, or parallel. These developmental trajectories should be interpreted under the condition that poor readers start at a lower level than non-poor readers. When non-poor readers improve their reading skills at faster pace than poor readers do, the development is divergent. This type of development process has been called the Matthew effect (Merton, 1968), also known as the rich getting richer and the poor get-

ting poorer. Bast and Reitsma (1998) found evidence for a Matthew effect. Morgan, Farkas and Hibel (2008) found evidence for a one-sided Matthew effect where the poor became poorer. The opposite is when poor readers improve their reading skills at a faster pace than non-poor readers do; in that case the development is convergent. Parrila, Aunola, Leskinen, Nurmi and Kirby (2005), and Aarnoutse, van Leeuwe, Voeten and Oud (2001) found support for a convergent development. Finally, when poor and non-poor readers develop at the same pace, then the development is parallel. Verhoeven and van Leeuwe (2003) and Aarnoutse and van Leeuwe (2000) found support for a parallel development. The common theme in all these studies is that they deal with the development across several grades. In contrast, we will study the development of poor and non-poor readers in grade 1.

The idea of a common instruction program together with an instruction program adapted for readers, average readers, and poor readers, is to minimize the number of pupils who fall behind and to keep all pupils motivated. The assignment of pupils to a specific category is very important for the individual pupil, because it will determine the amount and type of reading instruction. In the Learning to Read Safely programme the teacher decides to which category a pupil is assigned. At the beginning of the year, this decision is largely based on the performance on kindergarten tests and the readings skills. Juel (1988) and Smith (1997) found support for the stability of early classification. This finding paints a rather grim outlook for poor readers, as they have little chance on becoming a non-poor reader. However, not all scholars share this perspective. In a study covering 6 grades of primary school, Phillips, Norris, Osmond and Maynard, (2002) found that reading categories are more porous than might be concluded from other evidence (Juel, 1988; Smith, 1997). Lohman and Korb (2006) discuss several reasons why pupils are likely to be incorrectly classified, e.g. due to measurement errors, differential growth rates, and changes in the norming population.

5.1.4 Research questions

Reading development is commonly studied across several grades. We are, however, interested in the development of early reading, i.e. reading development in grade 1. During the first half of grade 1 reading instruction focuses on learning to decode words and to read short sentences, while in the second half of grade 1 the focus is on learning to read more complex words and on automatization of reading. Our first research question therefore is: How does word recognition develop during grade 1? Our second research question is: How does word recognition develop during grade 1, for pupils initially classified as poor and non-poor readers? There exists a controversy about the stability of the classification. Some scholars believe that pupils who are classified as poor readers always remain poor readers, or at least always belong to the lower percentiles of reading skills. Other scholars contradict this idea. Therefore, our third research question is: Are pupils tied to one category, i.e. poor readers or non-poor readers, during grade 1?

5.2 Methods

5.2.1 Participants

Our participants were 178 pupils from 8 Catholic schools located near the city of Utrecht in the Netherlands. The average age of the pupils at the start of grade 1 (August) was 6 years and 6 months (SD = 5.1 months). The sample consisted of 51% boys and 39% pupils from minority groups, mostly Turkish or Moroccan. The schools in our sample were asked to participate by college pupils who were doing their internships at those schools. All schools adopted the Learning to read safely programme (Mommers et al., 2003).

5.2.2 Measures

To study the development of the pupil's word recognition skills, we measured the speed of word recognition (WR). This is a measure of the pupil's ability to decode printed words (Aarnoutse & Kapinga, 2007). The pupil is presented a card with a list of 100 words of increasing difficulty. The unrelated words range from single syllable words to multi-syllabic words. The child is asked to read the words aloud as quickly and accurately as he or she can but without pressure. The test score is the number of words read correctly in 90 seconds. The test-retest correlation is .86 (Aarnoutse & Kapinga, 2007).

5.2.3 Procedure

We administered word recognition (WR) tests 4 times (November, January, March and May) in grade 1. The college pupils administered the WR-tests as part of their internship. They received special training sessions on how to administer these WR-tests. All WR-tests were administered with individual pupils in a separate and quiet room. All tests were done simultaneously (in the same week) on all 8 schools.

5.2.4 Data analyses

The data were analyzed in IBM statistics version 20. We used Repeated Measures ANOVA to analyse the development of word recognition of all pupils (N=178), as well as the development of poor readers versus non-poor readers. Pupils were classified as poor or non-poor readers using the WR-test of November. For the classification we used the percentile distribution in the manual (Aarnoutse & Kapinga, 2007). We choose the 20th percentile to create the groups of poor and non-poor readers. According to the manual the cut-off score for the November test is 15 words. Applying this cut-off score to our sample resulted in 18.5% poor readers (n=33) and 81.5% non-poor readers (n=145).

In order to analyse whether pupils are tied to one category during grade 1, we compared the classification in November with the classifications in January, March and May. Any disagreement in the consecutive classifications indicates that pupils are not tied to their classification, and thus indicates a classification error at January, March, or May compared to November. The classifications of January, March, or May are also based on a cut-off at the 20th percentile according to the manual; however, because pupils improve over time the actual cut-off scores will be higher. According to the manual, the cut-off score in January is 23 words, in March it is 35 words, and in May it is 44 words. If we apply these cut-off scores to our sample it results in 24.7% poor readers (n=44) in January, in 15.2% poor readers (n=27) in March, and in 19.1% poor readers (n=34) in May.

5.3 Results

Table 1 shows the univariate statistics of the WR tests for all pupils, as well as for the poor and non-poor readers. The average scores of the WR tests increase over time, for all pupils as well as for poor and non-poor readers. The biggest improvement occurs between January and March.

Table 1: Univariate statistics (mean (SD)) of the WR tests.

	All pupils (n=178)		Poor readers (n=33)		Non-poor readers (n=145)	
WR1 (November)	28.03	(16.89)	11.70	(2.35)	31.75	(16.57)
WR2 (January)	39.52	(19.59)	19.15	(6.27)	44.15	(18.60)
WR3 (March)	54.51	(19.75)	31.30	(10.29)	59.79	(17.45)
WR4 (May)	63.65	(20.70)	40.70	(16.32)	68.87	(17.86)

Figure 1 shows that the development of Word Recognition for all pupils (n=178) is approximately linear during grade 1. After each WR test (approximately 8 weeks in between), the scores increase roughly 10 words. A repeated measures ANOVA was performed to test whether there is a significant improvement over time. Mauchly's test indicated that the assumption of sphericity had been violated, CHI2(5)=197.02, p = .00, therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity (epsilon = .58). The results show that time has a significant effect on the WR test scores, F(1.74, 308.33) = 733.16, p = .00. We tested the significance of the improvement between successive measures using the contrast: repeated. The difference between November and January was significant (F(1, 177) = 308.73, p = .00), the difference between January and March was significant (F(1, 177) = 587.89, p = .00), and the difference between March and May was also significant (F(1, 177) = 249.37, p = .00).

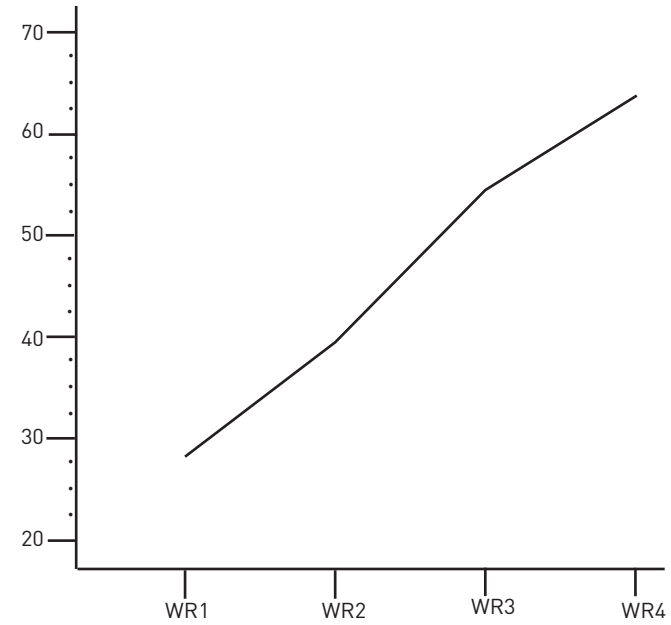


Figure 1: Development of word recognition for all pupils

Figure 2 shows that the development of Word Recognition for poor (n=33) and non-poor readers (n=145) is also approximately linear. After each WR test the number of words read in 90 seconds increases about 10 words. We performed a repeated measures ANOVA with poor/non-poor as a between groups factor, to test whether the development of poor and non-poor readers is convergent, divergent, or parallel. If the effect of time is moderated by poor/non-poor, then the development is either convergent or divergent. The factor poor/non-poor was created from the score on WR1, therefore the variable WR1 is not included in the analysis. Mauchly's test indicated that the assumption of sphericity had been violated, CHI2(5)=88.73, p = .00, therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity (epsilon = .72). The results show that time has a significant effect on the WR test scores, F(1.43, 251.84) = 315.88, p = .00. Furthermore, the effect of time is not moderated by poor/non-poor, F(1.43, 251.84) = 2.16, p = .13. This implies that the development is the same for poor and non-poor readers.

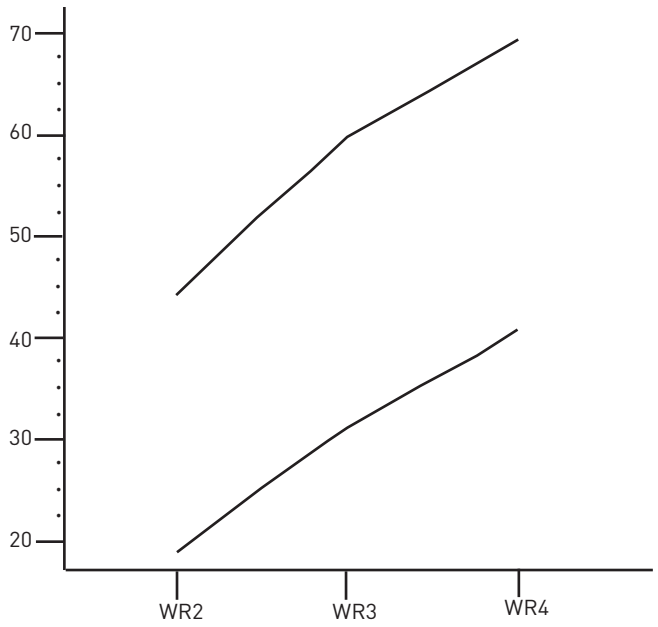


Figure 2: Development of word recognition for poor (bottom line) and non-poor readers (top line)

In the previous analyses pupils were tied to their classification during grade 1. In order to test whether they are tied, we have analyzed the transitions (over time) from poor to non-poor and the other way around. We created additional classifications for January, March, and May that indicate whether a pupil is a poor or a non-poor reader at that moment. The cut-off scores mentioned in the 'data analysis' section were used to determine who's a poor reader and who's not. Table 2 shows the transitions of poor to non-poor and of non-poor to poor between November and January, between November and March, and between November and May. If pupils would be tied to their classification, all off-diagonal cells would be empty. However, these off-diagonal cells are not empty indicating that pupils are not tied to their classification during the year. That is classification errors are made when the November classification is used in January, March, and May. It is possible to compute the percentage of classification errors by dividing the number of 'incorrectly' classified pupils by the total number of pupils. This results in 15.2% classification errors between WR1 and WR2, 11.2% between WR1 and WR3, and 16.3% between WR1 and WR4. We used the CHI2 test to test whether the observed frequencies deviate from the expected frequencies under the condition that no classification errors are made. However, because that would result in zero frequencies for the off-diagonal cells – resulting in a division by zero - we used an expected frequency of 1 for those cells. It follows that the number of classification errors from November to January is significant (CHI2(1) = 377.43, p < 00), from November to March is also significant (CHI2(1) = 185.46, p < .00), and from November to May is also significant (CHI2(1) = 372.49, p < .00).

Table 2: Transition of poor and non-poor readers from November to May.

	January		March		May	
November	Poor	Non-poor	Poor	Non-poor	Poor	Non-poor
Poor (n=33)	75.8%	24.2%	60.6%	39.4%	57.6%	42.4%
Non-poor (n=145)	13.1%	86.9%	4.8%	95.2%	10.3%	89.7%

5.4 Conclusions and Discussion

5.4.1 Conclusions

The development of word recognition in grade 1 is roughly linear. At average, pupils almost read twice as fast in May compared to November. This is quite a remarkable improvement.

We studied the development of word recognition for poor and non-poor readers. Pupils were classified as a poor reader when their score on word recognition was 15 or lower in November; other pupils were classified as non-poor readers. This cut-off value was taken from the manual (Aarnoutse & Kapinga, 2007) of the Word Recognition test and represented the 20th percentile. The results indicated that poor and non-poor readers develop at the same pace from January to May. We, therefore, conclude that poor and non-poor readers show a parallel development of word recognition. By definition, poor readers start at a lower level than non-poor readers. This paints a rather grim outlook for poor readers; they appear to be tied.

We studied whether pupils are indeed tied to their classification by inspecting the transitions between the November classification and successive classifications in January, March, and May. Based on the November classification 15.2% of the pupils are classified incorrectly in January, 11.2% are classified incorrectly in March, and 16.3% are classified incorrectly in May. This leads us to the conclusion that pupils are not tied to the classification made at the beginning of grade 1, i.e. after 10, weeks. About 40% of the poor readers in November manage to become non-poor readers in May, and about 10% of the non-poor readers in November become poor readers in May. The number of poor readers that become non-poor readers and the number of non-poor readers that become poor readers actually cancel each other out.

5.4.2 Discussion

There are many studies on the development of word recognition during primary school (e.g. Aarnoutse, Mommers, Smits, & van Leeuwe, 1986; Morgan, Farkas, & Hibel, 2008). To our knowledge our research is unique in that we study the development of reading in grade 1. At the beginning of grade 1, most pupils start with knowledge of the alphabet, but are not able to read. Our results show that after about 10 weeks, the average pupil reads almost 30 words in 90 seconds and 24 weeks later they read at average twice as fast. This is a remarkable achievement by both pupils and teachers. However, we all know that the average pupil does not exist and that pupils develop differently.

In our study we made a distinction between poor readers and non-poor readers. Poor readers can read at average almost 12 words in November (after about 10 weeks). Non-poor readers already read approximately 3 times as fast in the same period, almost 32 words. Despite this big difference, poor and non-poor readers develop at the same pace between January and May. The causes for this parallel development are, however, unclear. Given the large difference in reading skills in November it is odd that the difference doesn't increase more after November. One of the causes might be the instruction programme Learning to Read Safely that is used in the schools in our study. Teachers are trained to adapt their instruction to individual differences, with a special attention to poor readers. This extra attention might prevent the poor readers from further falling behind, although it could also result in poor readers to catch up with the non-poor readers. At the same time, the readers in the non-poor readers category are given extra difficult activities that should stimulate reading comprehension. This could potentially speed up reading development for the (non-poor) readers. This extra practice might result in accelerated development of word recognition of the (non-poor) readers. It seems that the sum of these processes results in a parallel development of word recognition between January and May for poor and non-poor readers.

In a study covering 6 grades of primary school, Phillips et al. (2002) found that reading categories are more porous than might be concluded from other evidence (Juel, 1988; Smith, 1997). We have found support for this finding, because we observed significant classification errors, i.e. pupils classified as poor readers in November have a large chance of becoming non-poor readers during the remainder of grade 1, while non-poor readers have a small chance of becoming poor readers in the same period. Lohman and Korb (2006) discuss several reasons for these classification errors. First there is the possibility that pupils develop their reading skills at a different pace, which among others could be attributed to the specific instructions that are part of the Learning to Read Safely programme, without ignoring the innate differences that exist between pupils. Another explanation that is often mentioned in the context of classification is regression towards the mean (Campbell & Stanley, 1963; Furby, 1973). This process refers to the effect of measurement error, or unreliability, in the word recognition scores. As a result one can expect that some pupils are incorrectly classified as poor readers, while they should have been classified as non-poor readers, and vice versa.

5.4.3 Limitations

The results of this study are that poor and non-poor readers in grade 1 develop at the same pace between January and May. This result is in agreement with the findings of Verhoeven and van Leeuwe (2003) and Aarnoutse and van Leeuwe (2000). Nevertheless, this result does not exclude the possibility that during the remainder of primary school there is a convergent development (Parrila, et al., 2005; Aarnoutse, et al. 2001), or a divergent development (Bast & Reitsma, 1998; Morgan, et al., 2008).

This study would have benefited from an early measure of word recognition, e.g. in the first week of grade 1. We could have used that early measure to classify pupils as poor or non-poor readers. This would enable us to study the development of poor and non-poor readers in grade 1 using 4 measures instead of 3. With the current measures we can only cover the development of poor and non-poor readers in the second half of grade 1. There is, however, a practical difficulty in using a WR-test at the very beginning of grade one, only very little pupils are able to read a few words at the beginning of grade 1. As a result, we would end up with a large group of pupils who are poor readers. Unfortunately, we cannot tell exactly what the results of a very early measure of word recognition would be, because we don't have one. We suggest including such a measure in a future study.

5.4.4. Practical implications

Our results indicated that pupils are not tied to the classification (poor/non-poor) made after about 10 weeks (November) in grade 1. We showed that at the end of grade 1, more than 40% of the poor readers have become non-poor readers, and more than 10% of the non-poor readers have become poor readers. A fair amount of pupils are crossing the boundaries from poor to non-poor readers and vice versa. When the teacher sticks to the initial classification during grade 1 this system doesn't work, because some pupils are not receiving the attention they deserve. In order to overcome this issue, it is important to measure word recognition skills frequently and to take appropriate action on the outcome. In the Netherlands, many schools are already adopting this idea. These schools write an action plan for a class every three months. This action plan is among other things based on observations in the classroom and tests that measure readings skills. Based on these observations and reading skill tests each pupil is classified as either an: excellent reader, average reader, or a poor reader. In the action plan goals are formulated for each group and also how to achieve those goals. At the end of every three-month period the achievements are evaluated and pupils are classified again as an excellent reader, average reader, or a poor reader. In addition, new goals are set. This way of repeatedly evaluating achievements and repeated adaptation of action plans prevents classifications errors, and thus to prevent that pupils receive instructions that are not adapted to individual differences.

Chapter 6

General Conclusions and Discussion

6.1 Introduction

In this study, which is concerned with the development of word recognition, the following overarching question was formulated: How does the word recognition of primary school pupils develop during the first grade with a Direct Instruction approach and with a Guided Co-Construction approach and what factors appear to influence the development of the children's word recognition? This overarching question encompassed the following three research questions: 1) Which instructional approach (DI or GCC) more effectively stimulates the development of word recognition in grade 1? 2) Which kindergarten pre-literacy skills appear to be important for the development of children's word recognition? 3) How does the word recognition of first grade pupils develop? To answer these questions we have examined the development of word recognition of 178 first grade pupils (in the ages of 6 and 7-year old), in a longitudinal study.

The following sections presents our answers to the research questions, their limitations, their implications for practice and finally, suggestions for further research.

6.2 Guided Co-Construction versus Direct Instruction

The focus of the first study (Chapter 2) was on the effect of two instructional approaches on the development of word recognition in grade 1: Direct Instruction (DI) and the Guided Co-Construction (GCC). For this study a field experiment with a pre-test/post-test control group design was undertaken. Two variants of the reading programme Learning to Read Safely were implemented among two different groups of pupils: a variant with a predominantly DI approach (i.e., the control group) and a variant with a more GCC approach (i.e., the experimental group). Our more specific question was: Is it 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.

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 further hypothesized that minority pupils could benefit more - than majority pupils - from direct instruction (DI).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).

The results of an overall analysis of the development of word recognition over time (i.e., throughout the whole 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 with better performance by the control group on the last measurement occasion. Majority pupils benefitted more from GCC but minority pupils more from DI. Minority pupils in the control group showed greatest progress.

6.3 Implementation

In order to substantiate the insights provided by the field experiment into the effectiveness of the different instructional approaches, the outcomes should be seen in the light of the question to what extent the instructional approaches were implemented in classroom practices. Therefore, in the second study (Chapter 3), the quality of the implementation of the two instructional approaches (DI and GCC) was assessed. The research question of this study was, whether the Direct Instruction and Guided Co-Construction approaches were implemented as intended. That is, did the activities of the pupils and teachers in the DI group show more characteristics of DI than the activities of the pupils and teachers in the GCC group? And conversely, did the activities of the pupils and teachers in the GCC group show more GCC characteristics than the activities of the pupils and teachers in the DI Group?

To answer these questions, we carefully described the implementation, collected systematic time-sampled observational data from the DI and GCC groups, and compared the degree of implementation for the two groups. The sums of the observations of teacher and pupil activities in the DI and GCC conditions were calculated and found to differ as expected. These qualitative and quantitative differences showed the implementation of the two instructional approaches to clearly differ in terms of actual classroom practice. Both the teacher and pupil activities reflect the differences which we expected to see for a Direct Instruction versus Guided Co-Construction approach to beginning reading instruction.

6.4 Predictors of word recognition

In Chapter 4 we were interested in the question what prerequisites influence the children's word recognition development? This third study was undertaken to see if children's word recognition could be predicted. The research question of this study was: Is there an effect of phonemic awareness, letter knowledge, and naming speed in kindergarten on word recognition after six month of reading instruction in grade1?

The results of this study showed that all of the kindergarten tests play a role in the speed of later word recognition. When these variables were subsequently entered into a hierarchical regression analysis and the nested nature of the cases controlled for by adding classes as fixed effects, the resulting model explained 53% of the variance in word recognition.

Letter knowledge and naming speed for digits exerted significant direct effects. In order, to quickly detect stagnation in the development of children's word recognition skills, it is thus recommend to measure letter knowledge, and naming speed for digits at the end of the kindergarten period. Such information enables us to to identify young children at risk for reading problems.

6.5 Development of word recognition

Finally in Chapter 5, we were interested in the development of word recognition of the group as a whole and two different sub-groups of readers namely: poor and non-poor readers. We asked ourselves the following three questions: How does word recognition develop during grade 1? How does word recognition develop during grade 1, for pupils initially classified as poor and non-poor readers? And are pupils tied to one category, i.e. poor reader or non-poor reader, during grade 1?

The results of this study showed that the development of word recognition in grade 1 is roughly linear and that poor and non-poor readers show a parallel development of word recognition. In addition, pupils are not tied to the classification made at the beginning of grade 1. About 40% of the poor readers in November manage to become non-poor readers in May, and about 10% of the non-poor readers in November become poor readers in May.

6.6 Limitations

The present studies are limited in certain ways. The first limitation concerns the lack of a random assignment of the schools to the control and the experimental conditions. That's why the research design became a quasi-experiment.

A second limitation concerns the use of the reading programme Learning to Read Safely in both the control and the experimental group. Although, some adaptations in the curriculum material were made, it should have been better to design a curriculum that is more in accordance with the instructional approach GCC.

A third limitation concerns the implementation of the two instructional approaches: GCC and DI. While we worked to ensure implementation fidelity by using carefully developed manuals and giving the teachers a solid training on the particular instructional approach, it is still possible that the implementation could be improved with the involvement of the teachers in the design of the curriculum, the development of the instructional approach and the design of the exercises. Not only would the teachers have a better understanding of the approach and the accompanying materials in such a manner; they might also have a stronger commitment to using and thus implementing the instructional approach. In addition, preparatory practice on the use of the instructional approach with pupils might also improve the implementation of the instructional approaches used in the present experiment. In other words, an innovation like GCC needs a long period of training and experience, otherwise teachers may fall back to their old patterns of teaching.

A last limitation is that this study would have benefited from an early measure of word recognition, e.g. in the first week of grade 1. We could have used that early measure as a pre-test in the first study (see Chapter 2) and to classify pupils as poor or non-poor readers in the fourth study (see Chapter 5) There is, however, a practical difficulty in using a WR-test at the very beginning of grade one, only very little pupils are able to read a few words at the beginning of grade 1. In the first study we solved this problem by using 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 measure important predictors of later word recognition (Aarnoutse, 2004; Aarnoutse, Van Leeuwe, & Verhoeven, 2000, 2005; Beernink, 2002; Verhagen, Aarnoutse, & Van Leeuwe, 2006, 2008). T-tests performed under the assumption of unequal variances showed the DI-group versus GCC-group do not differ on any pre-reading skills. Which indicate that both groups where initially comparable. In fourth study we solved this problem by using the first Word Recognition test (November) to classify pupils as poor or non-poor readers.

6.7 Implications for educational practice

A stagnation in the reading development for large groups of pupils seems not necessary, since research has shown that early reading ability has little or no relation with intelligence or social cultural background (Adams, 1990). However, many studies showed that about 10% of the pupils can be identified as 'functionally illiterate' (Sijtstra, van der Schoot, & Hemker, 2002; Wentink & Verhoeven, 2001). The question is, how can children best be stimulated in their reading development and how can stagnation be prevented as much as possible? The results of this dissertation generate five important implications for practice:

1. In the first study, we determined which approach was most effective for beginning reading, and whether pupils from minority versus majority socio-cultural back grounds might benefit differentially from different instructional approaches. The results of this study showed that GCC as a teaching-learning strategy is more effective in the classroom than DI. However, this effect faded during the second half of first grade. In addition, an interaction effect was found. Majority pupils benefited more from a GCC approach, while minority pupils profited more from DI. Therefore, from the first study it can be recommended to apply GCC as a teaching-learning strategy for beginning reading during the first half of the year. However, during this period, special attention should be given to minority pupils.

2. In the implementation study we carefully described the implementation of the DI and GCC approaches and compared the degree of implementation for the two

groups by systematic time-sampled observations of teachers and pupils activities from the DI and GCC groups. The sums of these observations were calculated and found to differ as expected. From this study it can be concluded that both instructional approaches: DI and GCC have been shown to be feasible for teaching beginning reading in real classroom settings.

3. The third study was undertaken to see if children's word recognition could be predicted with kindergarten tests. We examined the predictive value of phonemic awareness, letter knowledge, and naming speed in kindergarten on word recognition after six month of reading instruction. The results of a hierarchical regression analysis, showed that all kindergarten pre-literacy skills explained 53% of the variance in word recognition. Letter knowledge and naming speed for digits exerted significant direct effects. In order, to quickly detect stagnation in the development of children's word recognition skills, it is recommend to measure letter knowledge, and naming speed for digits at the end of the kindergarten period. Such information enables us to initiate early interventions to prevent further stagnation.

   Both The Central Institute for test development (Centraal Instituut voor Toetsontwikkeling: CITO) and Aarnoutse, Beernink and Verhagen (2008) developed a diagnostic program for teachers, namely, 'Language for pre-schoolers' (Taal voor kleuters) and 'Tests for Early Literacy' (Toetspakket voor Beginnende Geletterdheid). 'Language for pre-schoolers' is a programme that consists of two tests. A test for young pre-schoolers (pupils from 4 till 5 years old) and a test for older pre-schoolers (pupils from 5 till 6 years old). Both tests measure passive vocabulary and critical listening. The test for older pre-schoolers also measures phonological awareness, such as rhyme, phonemic analysis and phonemic synthesis and scripture orientation (general knowledge about books). The 'Tests for Early Literacy" entails tests for both young and older pre-schoolers. These tests measure naming speed, phonological awareness, letter knowledge and vocabulary. An important added value of 'Tests for Early Literacy' is the measuring of two important predictors of word recognition namely: naming speed and letter knowledge.

4. The results of the last study showed that the word recognition scores of all the pupils in grade 1 increase significantly. Apparently, the amount of instruction in reading influences the speeding up of the reading skills within this grade. Therefore, it is very important that there is sufficient time for beginning reading instruction. Every day, there should be enough time for all children to practice reading and on top of that additional time for poor readers.

5. In addition, the last study showed that children who were initially categorized as poor readers do not necessarily stay poor readers and unfortunately, children who were initially categorized as non-poor readers do not necessarily stay non-poor readers. When the teacher sticks to the initial classification during grade 1 some pupils are not receiving the attention they deserve. In order to overcome this issue, it is important to measure word recognition skills frequently (in grade 1 and the subsequent grades) and to take appropriate action on the outcome.

   In the Netherlands, many schools are already adopting this idea. These schools write an action plan for a class every three months. This action plan is among other things based on observations in the classroom and tests that measure readings skills. Based on these observations and reading skill tests each pupil is classified as either an: excellent reader, average reader, or a poor reader. In the action plan goals are formulated for each group and also how to achieve those goals. At the end of every three-month period the achievements are evaluated and pupils are classified again as an excellent reader, average reader, or a poor reader. In addition, new goals are set. This way of repeatedly evaluating achievements and repeated adaptation of action plans prevents classifications errors, and thus to prevent that pupils receive instructions that are not adapted to individual differences.

In conclusion, in the first half of the year, GCC as a teaching-learning strategy for beginning reading is feasible and more effective in the classroom than Direct Instruction. In addition, an interaction effect was found; majority pupils benefited more from a GCC approach, while minority pupils profited more from Direct Instruction. Furthermore, the word recognition scores of the pupils in grade 1 increase significantly. Therefore, it is very important that there is sufficient instruction time for beginning reading. Finally, in beginning reading, early detecting of stagnations is possible and recommendable. Pupils at risk sometimes catch up while others stay behind. Therefore it is very important to measure word recognition skills frequently. Only in this way, it is possible to follow accurately the development of word recognition, making targeted intervention possible.

6.7 Suggestions for further research

In our study we have chosen to measure the scores of word recognition as effect of two instructional approaches: DI and GCC. However, there could be more or other effects of these instructional approaches? For example, for dependent variables can also be taken other aspects of beginning reading like: reading comprehension, vocabulary or reading motivation. Or factors of more social processes of co-operation such as listening to others, trying to understand the perspective of others and being empathic and responsive to each other's needs (Ivey, 1994). Or with regard to specific verbal interactions that are related to learning such as giving explanations or providing elaborated help such as how to solve a problem or part of a problem (Webb, 1992). Therefore, for further research we recommend to measure not only the word recognition skills of children, but also the result of other important factors of reading and what can be called the processes of co-elaboration and co-construction.

ch6

A second important issue that needs to be mentioned is that word recognition can be distinguished into: speed and accuracy. These are not the same. Dyslexic children, for example, may recognize words in slow but precise manner or, alternatively, in a fast but inaccurate manner (Verhagen, 2009). The speed of word recognition is typically measured by having the pupil read as many words as possible from a list of unrelated words within a brief period of time, such as 90 seconds. Word recognition accuracy is typically measured using a graded word list without time pressure. In order to differentiate between children, moreover, the list used to measure the accuracy of word recognition should contain much more difficult words than the list used to measure the speed of word recognition. In sum, tests of which measure the speed of word recognition and the accuracy of word recognition measure partly different constructs. To detect reading problems, it is have been better to measure both the speed and accuracy of word recognition. For further research we recommend to measure both constructs (speed and accuracy) in a longitudinal study of children learning to read in languages with relative consistent orthographies (i.e. English or Dutch). In our study, we have chosen to measure only the speed of word recognition. We have made this decision because of two reasons: firstly, when children in the Netherlands are learning to read, learning to decode words accuracy is typically not the basic problem. In Dutch orthography, the speed of word recognition is the main problem (Verhagen, 2009). Secondly, because of practical reasons we have limited the measurement of reading tests in grade 1, after all in this grade the teachers already administer a lot of reading tests.

Finally, the four studies of this dissertation have been limited to the following two topics: (1) the development of children's word recognition during grade 1 and (2) the most effective instruction to stimulate this development. However, in order to improve reading instruction and to improve the development of children's word recognition further research is needed. Some interesting topics for future research could be: 'prevention of reading problems', 'reading motivation', 'effective reading programs' and 'effective teacher skills'.

Summary

The importance of reading in our time is evident. Someone who does not master this skill cannot function properly in our modern society. One of the main tasks of primary education is therefore to teach children to read. The purpose of beginning reading instruction is to help children master the many challenges of the written word including knowledge of the alphabetic system, the ability to decode new words and the ability to determine, integrate and recall the meanings of written words within texts.

This dissertation is concerned with effective reading instruction and the development of early literacy. Early literacy is the phase of literacy occurring after emerging literacy and before advanced literacy. Early literacy encompasses the development of reading from the beginning of kindergarten through the end of grade 1, with two goals standing central: discovery of the alphabetic principle and increased speed of word recognition (Verhoeven & Aarnoutse, 1999). Word recognition can be defined as the process of converting a sequence of letters into sounds for the identification of a word (Coltheart, 1978).

The development of beginning reading and especially word recognition has been widely studied both nationally and internationally. The following factors have been identified as important determinants of early word recognition: phonemic awareness; letter knowledge; rapid automatic naming of letters and digits (Aarnoutse, van Leeuwe, Voeten, & Oud, 2001; Bowers & Swanson, 1991; Hansen & Bowey, 1994; Näslund & Schneider, 1996; National Early Literacy Panel, 2008; Verhoeven & van Leeuwe, 2003; Wagner & Torgesen, 1987). This information, however, does not provide us with an answer to the question of which instructional approach is most suited to stimulate the development of children's early word recognition. Different instructional approaches can be expected to produce different results. And one of the major questions to be addressed — and the starting point for this dissertation — is whether knowledge should be provided by teachers or generated by the learners themselves (Rosenshine, Meister, & Chapman, 1996). Effective instruction would seem to call for a third alternative to overcome this dilemma. This third alternative can be found in the guided co-construction of knowledge (Hardman, 2008; Mercer, 1995; Terwel, van Oers, van Dijk & van den Eeden, 2009; van Schaik, Terwel, & van Oers, 2014). Guided co-construction (GCC) as an instructional approach requires the following core elements: an explicit role for the teacher in whole-class instruction, pupil guidance and the scaffolding of pupil learning; cooperative learning on the part of pupils; and the construction of insight, skill and behaviour by pupils on the basis of their accumulated knowledge and experiences.

In this dissertation, the effects of two instructional approaches on the development of young children's word recognition were compared: a providing approach Direct Instruction (DI) versus an instructional approach which included a cooperative learning component (GCC). DI is a approach for teaching which utilizes carefully planned lessons designed around small learning increments. The instruction is highly structured and describes or even scripts classroom activities in considerable detail. DI has been shown to be an effective instructional approach for teaching children and particularly children at risk for reading difficulties to read (Adams, 1990; Bus & van IJzendoorn, 1999; Chall, 1996; Ehri, Nunes, Stahl, & Willows, 2001; Hattie, 2008). However, in recent reviews and meta-analyses, Raudenbush (2009) and Slavin, Lake, Chambers, Cheung and Davis (2009) have found other instructional approaches which include a cooperative learning component to be particularly effective for the teaching of early word recognition.

In light of the above, the overarching question for the studies presented in this dissertation became: How does the word recognition of primary school pupils develop during the first grade with a Direct Instruction approach and with a Guided Co-Construction approach and what factors appear to influence the development of the children's word recognition? This overarching question encompassed the following three research questions:

1. Which instructional approach (DI or GCC) more effectively stimulates the development of word recognition in grade 1?
2. Which kindergarten pre-literacy skills appear to be important for the development of children's word recognition?
3. How does the word recognition of first grade pupils develop?

To answer these questions, the development of the word recognition of 178 first grade pupils was examined, in a longitudinal study.

Guided Co-Construction versus Direct Instruction

In the first study (Chapter 2), we determined which form of instruction best stimulates the development of children's early word recognition. A field experiment with a pre-test/post-test control group design was undertaken. Two instructional variants of the reading programme Learning to Read Safely were implemented with separate groups of pupils: DI with the control group and GCC with the experimental group. Teacher training was provided on the different instructional approaches and special materials were developed and supplied to facilitate either the DI or GCC.

The development of the 178 children's word recognition throughout the course of grade 1 was compared to determine which instructional approach was most effective. More specifically, it was asked: Is it 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.

Drawing on the outcomes of an earlier series of studies (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 of those who received DI. It was further hypothesized that the pupils would benefit differently from the two instructional approaches: Minority pupils would benefit more from direct instruction (DI) but majority pupils more from GCC (Leseman & de Jong, 1998).

The results of a repeated measures analysis showed the pupils in the GCC group to outperform the pupils in the DI group. However, these results also showed the differences between the groups to gradually diminish during the course of grade 1 and the pupils in the DI group to almost catch up to the pupils in the GCC group by the end of the school year. In addition, a significant interaction between instructional condition and the socio-cultural background of the pupils was found: The majority pupils indeed benefitted more from GCC while the minority pupils benefitted more from DI. The minority pupils in the control group were found to show the most progress, and closer inspection of the development of the word recognition of the 109 majority pupils separately showed their performance differences between the conditions to be relatively constant throughout the year.

Implementation

In the second study (Chapter 3), the quality of the implementation of the two instructional approaches — DI and GCC — was assessed. The specific research question was whether the two instructional approaches were implemented as intended. That is, did the activities of the pupils and teachers in the DI group show more characteristics of DI than the activities of the pupils and teachers in the GCC group? And conversely, did the activities of the pupils and teachers in the GCC group show more GCC characteristics than the activities of the pupils and teachers in the DI Group? To answer the research question, we carefully described the implementation process for the two instructional; systematically collected time-sampled observational data from the DI and GCC groups; and compared the quality of implementation for the two groups. The sums of the systematic time-sampled observations for the pupil and teacher activities in the DI and GCC conditions were then calculated and found to differ as expected: The instructional approaches were implemented as intended.

Prediction of word recognition

The third study (Chapter 4) was undertaken to examine the prediction of the children's word recognition. Among the cognitive prerequisites for learning to read (Adams, 1990; Verhagen, 2009) are phonemic awareness (Aarnoutse, 2004), letter knowledge as a basis for understanding the alphabetic principle (Bowey, 2005) and naming speed (Verhagen, 2009). The specific research question in this third study was therefore: Is there an effect of phonemic awareness, letter knowledge and naming speed in kindergarten on children's word recognition after six month of reading instruction in grade 1?

The results showed all of the kindergarten pre-reading skills to indeed play a role in the children's later speed of word recognition. When the background variables were subsequently entered into a hierarchical regression analysis and the nested nature of the cases controlled for with the addition of 'classes' as fixed effects, the resulting model explained 53% of the variance in word recognition. Letter knowledge and naming speed for digits in kindergarten showed significant direct effects on the children's word recognition, moreover.

In order, to quickly detect stagnation in the development of children's word recognition skills, it is recommend to measure letter knowledge, and naming speed for digits at the end of the kindergarten period. Such information enables us to to identify young children at risk for reading problems

Development of word recognition

In the fourth and final study (Chapter 5), the development of the children's word recognition during the course of first grade was examined. Children with different levels of ability can be expected to show different patterns of reading performance over time (e.g., Stanovich, 1986; Williamson, Appelbaum, & Enpanchin, 1991). The specific research questions in this fourth study were as follows: How does word recognition develop during grade 1? How does word recognition develop during grade 1, for pupils initially classified as poor versus non-poor readers? And are pupils tied to one category, i.e. poor reader or non-poor reader, during grade 1?

The results of this study showed the development of word recognition in grade 1 to be roughly linear and the poor and non-poor readers to show parallel patterns of development in their word recognition. Nevertheless, the pupils did not remain tied to their initial classification as a 'poor' or 'non-poor' reader at the beginning of grade 1: About 40% of the poor readers identified in November managed to become non-poor readers by May of the same school year while about 10% of the non-poor readers identified in November had become poor readers in May.

Suggestions for future research

In this dissertation, word recognition was chosen as the most important variable for assessing the effects of alternative approaches to early reading instruction (i.e., DI versus GCC). Other variables — such as vocabulary, reading comprehension and reading motivation — should also be examined in future research. In addition, variables such as listening to others and trying to understand the perspectives of others might be examined as also relevant to assess the effects of DI and GCC (Ivey, 1994).

The speed and accuracy of word recognition are commonly distinguished (Verhagen, 2009). In the studies presented in this dissertation, only the speed of the children's word recognition was measured. This was done because speed of recognition has been found to be the main problem encountered by children having difficulties learning to read. In future research, both the speed and accuracy of children's word recognition should nevertheless be examined as these are known to interact and relate differently over time (Verhagen, 2009).

Finally, additional research is needed to improve reading instruction and further stimulate the development of children's word recognition. Among the topics which should be addressed are the prevention of reading problems, the fostering of reading motivation, the implementation of particularly effective reading programmes and effective teacher skills.

Implications for educational practice

The results of this dissertation have some important implications for both the stimulation of reading development and prevention of stagnation. The results of the first two studies showed both DI and GCC to be feasible approaches for the teaching of beginning reading in real classroom settings. Overall, GCC was found to be most effective for the stimulation of word recognition. However, this effect faded during the second half of first grade. In addition, an interaction effect was found. Majority pupils benefited more from a GCC approach, while minority pupils profited more from DI. Therefore, it can be recommended to apply GCC as a teaching-learning strategy for beginning reading during the first half of the year. However, during this period, special attention should be given to minority pupils.

Also on the basis of the present results, it is recommended that educational practice consider measurement of letter knowledge and naming speed for digits at the end of the kindergarten period. Such measurement can alert us to stagnation in the development of a child's early word recognition skills and allow us to initiate early intervention aimed at the prevention of further delay.

The present results show the devotion of sufficient time to reading instruction during grade 1 is important. Additional time for pupils identified as poor readers is also beneficial and therefore called for.

To conclude, the fourth and final study in this dissertation showed those children who were initially categorized as poor readers to not necessarily remain poor readers; unfortunately, some 10% of the children initially categorized as non-poor readers did not remain this and had thus become poor readers by the end of first grade. When a teacher sticks to the initial classification of a child at the start of grade 1 for the remainder of the school year, some pupils will thus not received the attention which they need and deserve. In order to overcome this obstacle, it is thus important that word recognition skills be frequently measured and appropriate action be taken depending on the outcome. Frequent measurement is important in not just grade 1 but also in the subsequent grades of primary school.

References

Aarnoutse, C. (2004). *Ontwikkeling van beginnende geletterdheid* [Development of early literacy]. Nijmegen, The Netherlands: Radboud Universiteit Nijmegen.

Aarnoutse, C., Beernink J., &Verhagen, W. (2010). *Toetspakket beginnende geletterdheid* [Beginning literacy test package]. Amersfoort, The Netherlands: CPS.

Aarnoutse, C., & Kapinga, T. (2007). *Technisch lezen 3* [Technical reading 3]. Ridderkerk, The Netherlands: Onderwijs Advisering.

Aarnoutse, C. A. J., Mommers, M. J. C., Smits, B. W. G. M., & Van Leeuwe, J. F. J. (1986). De ontwikkeling en samenhang van technisch lezen, begrijpend lezen en spellen [The development and connection of decoding, reading comprehension and spelling]. *Pedagogische Studiën, 63*, 97-108.

Aarnoutse, C., & Van Leeuwe, J. (2000). Development of poor and better readers during the elementary school. *Educational Research and Evaluation, 6,* 251-278. doi:10.1076/1380-3611(200009)6:3;1-A;FT251

Aarnoutse, C., Van Leeuwe, J., & Verhoeven, L. (2000). De ontwikkeling van beginnend geletterdheid [Development of early literacy]. *Pedagogische Studiën, 77,* 307-325.

Aarnoutse, C., Van Leeuwe, J., & Verhoeven, L. (2005). Early literacy from a longitudinal perspective. *Educational Research and Evaluation, 11,* 253-275. doi:10.1080/08993400500101054

Aarnoutse, C., Van Leeuwe, J., Voeten, M., & Oud, H. (2001). Development of decoding, reading comprehension, vocabulary and spelling during the elementary school years. *Reading and Writing: An Interdisciplinary Journal, 14,* 61–89. doi: 10.1023/A:1008128417862

Aarnoutse, C., & Verhagen, W. (2001). *Fonemische synthesetoets* [Phonemic synthesis test]. Unpublished manuscript.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge: Massachusetts Institute of Technology.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the commission on reading.* Urbana: University of Illinois.

Barnes, D. R., & Todd, F. (1995). *Communication and learning revisited: Making meaning through talk.* Portsmouth, NH: Heinemann.

Bast, J., & Reitsma, P. (1998). Analyzing the development of individual difference in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology, 34,* 1373-1399. doi:10.1037/0012-1649.34.6.1373

Beernink, J. (2002). *Op weg naar geletterdheid* [On the way to literacy] (Unpublished master's thesis). Radboud University Nijmegen, The Netherlands.

Bennett, N. (1987). Cooperative learning: Children do it in groups – or do they? *Educational and Child Psychology, 4*(3-4), 7-18.

Bennett, N., Jordan, J., Long, G., & Wade, B. (1976). *Teaching styles and pupil progress.* Cambridge, MA: Harvard University Press.

Biemans, H. A., Jongmans, C. T., De Jong, F. P. C. M., & Bergen, T. C. M. (1999). Perceptions of teachers' instructional behavior in secondary agricultural education. *Journal of Agricultural Education and Extension, 5,* 231-238. doi:10.1080/ 13892249985300031

Blachman, B. A. (2000). Phonological awareness. In M. L. Kamil, P. B., Mosenthal, P. D.
Pearson, & R. Barr (Eds.), *Handbook of Reading Research* (Vol. 3, pp. 483-502). Mahwah, NJ: Erlbaum.

Bond, G. L., & Dijkstra, R. (1997). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 32,* 342-345. doi:10.1598?RRQ.32.4.4

Bowers, P. G. (1997). *Exploration of the link between processes tapped by rapid naming tests and orthographic skill.* Paper presented at the Society for the Scientific Study for Reading, Chicago, IL.

Bowers, P. G., Golden, J., Kennedy, A., & Young, A. (1994). Limits upon orthographic knowledge due to processes indexed by naming speed. In V. W. Berninger (Ed.), *The varieties of orthographic knowledge I: Theoretical and developmental issues* (pp. 173–218). Dordrecht/Boston/London: Kluwer Academic Press.

Bowers, P. G., & Swanson, L. B. (1991). Naming speed deficits in reading disability: Multiple measures of a single process. *Journal of Experimental Child Psychology, 51,* 195-219. doi:10.1016/0022-0965(91)900332-N

Bowey, J. A. (2005). Predicting individual differences in learning to read. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading* (pp. 155-173). Oxford, UK: Blackwell.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, Learning and Instruction. Essays in Honor of Robert Glaser* (pp. 393-451). Hillsdale, NJ: Lawrence Erlbaum.

Bruner, J. (1985). Vygotsky: A historical and conceptual perspective. In J. V. Wertsch (Ed), *Culture, Communication and cognition: Vygotskian perspectives.* Cambridge, UK: Cambridge University Press.

Bus, A. G., & Van IJzendoorn, M. H. (1999). Phonological awareness and early reading: a meta-analysis of experimental training studies. *Journal of Education Psychology, 91,* 403-414. doi:10.1037//0022-0663.91.3.403

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand-McNally.

Chall, J. S. (1996). *Learning to read: The great debate* (3rd ed.). New York, NY: McGraw-Hill.

Coles, G. (2000). *Misreading reading: The bad science that hurts children.* Portsmouth, NH: Heinemann.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 135-155). New York, NY: Academic Press.

Creemers, B. P. M. (1991). *Effectieve instructie: Een empirische bijdrage aan de verbetering van het onderwijs in de klas.* [Effective instruction: An empirical contribution to improvement of education in the class]. 's-Gravenhage, The Netherlands: SVO.

Cutting, L. E., & Denckla, M. B., (2001). The relationship of rapid serial naming and word reading in normally developing readers: An exploratory model. *Reading and Writing: An Interdisciplinary Journal, 14,* 673-705. doi:10.1023/A:1012047622541

De Jager, B. (2002). *Teaching reading comprehension: The effects of direct instruction and cognitive apprenticeship on comprehension skills and metacognition.* Groningen, The Netherlands: Rijksuniversiteit Groningen.

De Jong, P. F., & Leseman, P. P. M. (2001). Lasting effects of home literacy on reading achievement in school. *Journal of School Psychology, 39,* 389–414. doi:10.1016/ S0022-4405(01)00080-2

De Jong, P. F., Van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91,* 450–476. doi:10.1037//0022-0663.91.3.450

De Jong, P. F., Van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology, 95,* 22-40. doi:10.1037//0022-0663.95.1.22

Dewey, J. (1943). *The School and Society.* Chicago, IL: University Press.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38,* 78-103. doi:10.1598/ RRQ.38.1.4

Durlak, J. A. (2010). The importance of doing well in whatever you do: a commentary on the special section 'Implementation research in early childhood education'. *Early Childhood Research Quarterly, 25,* 348–357. doi:10.1016/j.ecresq.2010.03.003

Durlak, J. A., & Dupre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting the implementation. *American Journal of Community Psychology, 41,* 327-350. doi:10.1007/s10464-008-9165-0

Edmonds, R. (1977). Effective schools for the urban poor. *Educational Leadership, 37,* 15-24.

Ehri, L. C. (1991). Development of the ability to read words. In R. Barr, M. L. Kamil, P. Mosenthal,

& P. D. Pearson (Eds.), Handbook of Reading Research. New York, NY: Longman.
Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71,* 393-447. doi:10.3102/ 00346543071003393
Ehri, L. C., & Sweet, J. (1991). Fingerpoint research of memorized text: What enables beginners to process the print? *Reading Research Quarterly, 26,* 442-462.
Ehri, L. C., & Wilce, L. S. (1987). Does learning to spell help beginners learn to read words? *Reading Research Quarterly, 18,* 47-65. doi:10.2307/747720
Evans, M. A., & Carr, T. H. (1985). Cognitive abilities, conditions of learning, and the early development of reading skill. *Reading Research Quarterly, 20,* 327-350. doi:10. 2307/748022
Fleiss, J. L. (1981). *Statistical methods for rates and proportions.* New York, NY: John Wiley.
Freudenthal, H. (1991). *Revisiting mathematics education.* Dordrecht, Boston, London: Kluwer Academic Press.
Fullan, M. (1991). *The new meaning of educational change.* London, UK: Cassell.
Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Education Research, 47,* 355-397. doi:10.3102/00346543047002335
Furby, L. (1973). Interpreting regression toward the mean in developmental research. *Developmental Psychology, 8,* 172-179. doi:10.1037/h0034145
Gentry, J. R. (2006). *Breaking the code. The New Science of beginning reading and writing.* Portsmouth, NH: Heinemann.
Goodlad, J. I. (1994). Curriculum as a field of study. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (pp. 1262-1267). Oxford, UK: Pergamon Press.
Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention re¬search: Do we really know how treatments are implemented? *Learning Disabilities Research and Practice, 15,* 198–205.
Hansen, J., & Bowey, J. A. (1994). Phonological analysis skills, verbal working memory, and reading ability in second-grade children. *Child Development, 65,* 938-950. doi:10. 2307/1131429
Hardman, F. (2008). The guided co-construction of knowledge. In M. Martin-Jones, A. De Mejia, & N. Hornbeger (Eds.), *Encyclopaedia of language and education: Vol. 3. Discourse and education* (pp. 253-264). New York, NY: Springer.
Hattie, J. (2008). *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement.* London, UK: Routledge.
Hecht, S. A., Burgess, S. R., Torgesen, J. K., Wagner R. K., & Rashotte C. A. (2000). Explaining social class differences in growth of reading skills from beginning Kindergarten through fourth-grade: The role of phonological awareness, rate of access and print knowledge. *Reading and Writing: An Interdisciplinary Journal, 12,* 99-127. doi:10.1023/A:1008033824385
Henneman, K., Kleijnen, M., & Smits, A. E. H. (2004). Protocol dyslexie VO. *Signalering, diagnose en begeleiding* [Protocol dyslexia secondary education. Signaling, diagnosis and counseling]. 's Hertogenbosch, The Netherlands: KPC.
Hoek, D. J. (1998). *Social and cognitive strategies in co-operative groups. Effects of strategy instruction in secondary mathematics.* Amsterdam, The Netherlands: Universiteit van Amsterdam.
Hoek, D., Terwel, J., & Van den Eeden, P. (1997). Effects of training in the use of social and cognitive strategies: An intervention study in secondary mathematics in co-operative groups. *Educational Research and Evaluation, 3,* 364–389. doi:10.1080/ 138036197030405
Hoek, D. J., Terwel, J., Van Hout-Wolters, B. H. A. M., & Van den Eeden, P. J. W. M. (2003). Effects of quality of implementation on the learning outcomes of secondary education mathematics curriculum. *Curriculum and Teaching, 18,* 5-25. doi:10.7459/ ct/18.1.02
Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology, 65,* 599-610. doi:10.1037//0022-006X.65.4.599
Houtveen, A. A. M., Brokamp, S. K., & Smits, A. E. H. (2012). *Lezen, lezen, lezen! Achtergrond en evaluatie van het LeesInterventie-project voor Scholen met een Totaalaanpak (LIST)* [Read, read, read! Background and evaluation of the reading intervention project for schools with a total approach (LIST)]. Utrecht, The Netherlands: Hogeschool Utrecht, Kenniscentrum Educatie van de Faculteit Educatie.
Inspectie van het Onderwijs. (2006). *De staat van het onderwijs. Onderwijsverslag 2004/2005* [The state of education. Education Report 2004/2005]. Utrecht, The Netherlands: Inspectie van het Onderwijs.
Inspectie van het Onderwijs. (2008). *Basisvaardigheden taal in het basisonderwijs. Een onderzoek naar het niveau van de taalvaardigheden in het basisonderwijs en naar verschillen tussen scholen met lage, gemiddelde en goede taalresultaten* [Skills of language in primary education. A research into the level of language skills in primary education and disparities between schools with low, average and good language performance]. Utrecht, The Netherlands: Inspectie van het Onderwijs.
Ivey, A. (1994). *Intentional interviewing and counseling* (3rd ed.). Pacific Grove, CA: Brooks/Cole.
Jacobson, C. (1999). How persistent is reading disability? Individual growth curves in reading. *Dyslexia, 5,* 78-93. doi:10.1002/(SICI)1099-0909(199906)5:2<78:: AIDDYS127>3.0.CO;2-8
Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 78,* 243–245. doi:10.1037// 0022-0663.80.4.437
Kame'enui, E.J. & Carnine, D.W. (Eds.). (1998). *Effective teaching strategies that accommodate diverse learners.* Englewood Cliffs, NJ: Prentice Hall.
Keijzer, R. (2003). *Teaching formal mathematics in primary education. Fraction learning as mathematizing process.* Amsterdam, The Netherlands: Vrije Universiteit van Amsterdam.
Knight, P. (2001). Complexity and curriculum: A process approach to curriculum-making. *Teaching in Higher Education, 6,* 369-381. doi:10.1080/13562510120061223
Lansink, N. (2009). Taal voor kleuters [Language for pre-schoolers]. Arnhem, The Netherlands: Citogroep.
Lendrum, A., & Humphrey, N. (2012). The importance of studying the implementation of interventions in school settings. *Oxford Review of Education, 38,* 635-652. doi:10.1080/03054985.2012.734800
Leppänen, U., Niemi, P., Aunola, K., & Nurmi, J. E. (2004). Development of reading skills among preschool and primary school pupils. *Reading Research Quarterly, 39,* 72-93. doi:10.1598/RRQ.39.1.5
Leseman, P. P. M., & De Jong, P. F. (1998). Home literacy: Opportunity, instruction, cooperation, and social–emotional quality predicting early reading achievement. *Reading Research Quarterly, 33,* 294–318. doi:10.1598/RRQ.33.3.3
Lohman, D. F., & Korb, K. (2006). Gifted today but not tomorrow? Longitudinal changes in ability and achievement during elementary school. *Journal for the Education of the Gifted, 29,* 451-484. doi:10.4219/jeg-2006-245
Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. *Developmental Psychology, 36,* 596-613. doi:10.1037//0012-1649.36.5.596
Magliaro, S. G., Lockee, B. B., & Burton, J. K. (2005). Direct instruction revisited: A key model for instructional technology. *Educational Technology Research and Development, 53*(4), 41-55. doi:10.1007/BF02504684
Manis, F. R., Doi, L. M., & Bhada, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. *Journal of Learning Disabilities, 33,* 325–333. doi:10.1177/002221940003300405

McCaslin, M., & Burross, H. L. (2011). Research on individual differences within a sociocultural perspective: Co-regulation and adaptive learning. *The Teachers College Record, 113,* 325-349.
McGill-Franzen, A., Ward, N., Goatley, V., & Machado, V. (2002). Teachers' use of new standards, frameworks, and assessments: Local cases of NYS elemen¬tary grade teachers. *Reading Research & Instruction, 4,* 127-148. doi:10.1080/ 19388070209558362
Mercer, N. (1995). *The guided construction of knowledge: Talk among teachers and learners.* Clevedon, UK: Multilingual Matters.
Merton, K. (1968). The Matthew effect in science. *Science, 159*(3810), 56–63.
Ministry of Education, Culture and Science. (2006). *Kerndoelen primair onderwijs* [Core learning goals Primary education]. Den Haag, The Netherlands: Delta Hage.
Mommers, M., Verhoeven, L., Koekebacker, E., Van der Linden, S., Stegeman, W., & Warnaar, J. (2003). *Veilig leren lezen* [Learning to read safely]. Tilburg, The Netherlands: Zwijsen.
Morgan, P. L., Farkas, G., & Hibel, J. (2008). Matthew effects for whom? *Learning Disability Quarterly, 31*, 187-198.
Näslund, J. C., & Schneider, W. (1996). Kindergarten letter knowledge, phonological skills, and memory processes: Relative effects on early literacy. *Journal of Experimental Child Psychology, 62,* 30-59. doi:10.1006/jecp.1996.0021
National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel.* Washington, DC: National Institute for Literacy. Retrieved from htt://www.nifl.gov/earlychildhood/NELP/NELPreport.html
National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: Government Printing Office.
Parrila, R., Aunola, K., Leskinen, E., Nurmi, J. E., & Kirby, J. R. (2005). Development of individual differences in reading: Results from longitudinal studies in English and Finnish. *Journal of Educational Psychology, 97,* 299-319. doi:10.1037/0022-0663.97. 3.299
Phillips, L. M., Norris, S. P., Osmond, W. C., & Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first through sixth grades. *Journal of Educational Psychology, 94,* 3–13. doi:10.1037//0022-0663.94.1.3
Popp, H. M., & Lieberman, M. (1977). *A study of the relationship of student achievement to components of reading programs and environmental characteristics* [Final Report No. 400-75-0064]. Cambridge, MA: Harvard Graduate School of Education.
Pressley, M., Graham, S., & Harris, K. (2006). The state of educational interventions research as viewed through the lens of literacy intervention. *British Journal of Educational Psychology, 76,* 1-19. doi:10.1348/000709905X66035
Prinsen, F. R. (2007). XS2KNOW *(Access to knowledge). Differences in participation and learning outcomes of students in computer supported collaborative learning.* Amsterdam, The Netherlands: Vrije Universiteit van Amsterdam.
Prinsen, F. R., Terwel, J., Zijlstra, B. J. H., & Volman, M. L. L. (2013). The effects of guided elaboration in a CSCL programme on the learning outcomes of primary school students from Dutch and immigrant families. *Educational Research and Evaluation, 2013, 19*, 39–57. doi:10.1080/13803611.2012.744694
Raudenbush, S. W. (2009). The *Brown* legacy and the O'Connor challenge: Transforming schools in the images of children's potential. *Educational Researcher, 38,* 169-180. doi:10.3102/0013189X09334840
Revicki, D. A. (1981). Models for measuring programme implementation, a review and critique. *Journal of curriculum studies, 12,*65-68.
Rosenshine, B. (1979). Content, time, and direct instruction. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings and implications* (pp. 28-56). Berkeley, Cal: McCutcheon.
Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66,* 181-221. doi:10.3102/00346543066002181
Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376-391). New York, NY: MacMillan.
Ryder, R. J., Burton, J. L., & Silberg, A. (2006). Longitudinal study of direct instruction effects from first through third grades. *The Journal of Educational Research, 99,* 179-191. doi:10.3200/JOER.99.3.179-192
Sawyer, R. K. (2006). Introduction: The new science of learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 1-16). New York, NY: Cambridge University Press.
Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, A. J. Capute, & B. Shapiro (Eds.), *Specific reading disability: A view of the spectrum* (pp. 243-274). Hillsdale, NJ: Lawrence Erlbaum Associates.
Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. Educational Researcher, 27(2), 4-13. doi:10.3102/013189X027002004
Share, D., Jorm, A., Maclean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology, 76,* 1309-1324. doi:10.1037//0022-0663.76.6.1309
Sijtstra, J., Van der Schoot, F., & Hemker, B. (2002). *Balans van het taalonderwijs aan het einde van de basisschool 3. Uitkomsten van de derde peiling in 1998* [Balance of language education at the end of the elementary school 3. Results of the third poll in 1998]. Arnhem, The Netherlands: Citogroep.
Simmons, D. C., & Kame'enui, E. J. (Eds.). (1998). *What reading research tells us about children with diverse learning needs. Bases and basics.* Mahwah, NJ: Lawrence Erlbaum Associates.
Slavin, R. E. (1990). *Cooperative learning: Theory, research and practice.* Englewood-Cliffs, NJ: Prentice Hall.
Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research, 79,* 1391-1466. doi:10.3102/0034654309341374
Smith, S. S. (1997). A longitudinal study: The literacy development of 57 children. In C. Kinzer, K. A. Hinchman, & D. J. Leu (Eds.), *Inquiries in literacy theory and practice* (pp. 250–264). Chicago, IL: National Reading Conference.
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,* 353–374. doi:10.1080/13803611.2012. 682803
Snyder, J., Bolin, F., & Zumwalt, K. (1992). Curriculum implementation. In P. Jackson (Ed.), *Handbook of research on curriculum* (pp. 402-435). New York, NY: MacMillan.
Spring, C., & Capps, C. (1974). Encoding speed, rehearsal, and probed recall of dyslexic boys. *Journal of Educational Psychology, 66,* 780 – 786. doi:10.1037/h0037472
Spring, C., & Davis, J. M. (1988). Relations of digit naming speed with three components of reading. *Applied Psycholinguistics, 9,* 315–334. doi:10.1017/S0142716400008031
Stahl, S. A. (1999). Why innovations come and go (and mostly go): The case of whole language. *Educational Researcher, 28*(8), 13-22. doi:10.3102/0013189X028008013
Stahl, S., & Miller, P. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research, 59,* 87-116. doi:10.3102/00346543059001087
Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 22,* 360-406. doi:10.1598/RRQ.21.4.1

Stoep, J. (2008). *Opportunities for early literacy development: Evidence for home and school support.* Apeldoorn, The Netherlands: Garant Uitgevers nv.
Sunseth, K., & Bowers, P. G. (1997). *The relationship between digit naming speed and orthography in children with and without phonological deficits.* Paper presented at the annual conference for the Society for the Scientific Study of Reading, Chicago, IL.
Terwel, J., Van Oers, B., Van Dijk, I. M. A. W., & Van den Eeden, P. (2009). Are representations to be provided or generated in primary mathematics education? Effects on transfer. *Educational Research and Evaluation, 15,* 25-44. doi:10.1080/13803610802481265
Van den Bos, K. P., & Lutje Spelberg, H. C. (2010). *Continu benoemen en woorden lezen* [Continuous naming and word reading]. Groningen, The Netherlands: Boom Testuitgevers.
Van den Bos, K. P., Zijlstra, B. J. H., & Lutje Spelberg, H. C. (2002). Life-span data on continuous naming speed of numbers, letters, colors, and pictured objects, and word reading speed. *Scientific Studies of Reading, 6,* 25-49. doi:10.1207/ S1532799XSSR0601_02
Van Dijk, I. M. A. W. (2002). *The learner as designer: Processes and effects of an experimental programme in primary mathematics education.* Amsterdam, The Netherlands: VU University Amsterdam.
Van Dijk, I. M. A. W., Van Oers, H. J. M., & Terwel, J. (2003). Providing or designing? Constructing models in primary maths education. *Learning and Instruction, 13,* 53-72. doi:10.1016/S0959-4752(01)00037-8
Van Schaik, M. (2010). *Co-constructing models as tools in vocational practice. Learning in a knowledge-rich environment.* Zoetermeer, The Netherlands: Free Musketeers.
Vellutino, F., & Fletcher, J. (2005). Developmental dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 362–378). Malden, MA: Blackwell.
Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88,* 601–638. doi:10.1037//0022-0663.88.4.601
Verhagen, W. (2009). *Predicting early word recognition and spelling.* Nijmegen, The Netherlands: Radboud University Nijmegen.
Verhagen, W., & Aarnoutse, C. (2001). *Fonemische analyse en synthese van pseudowoorden* [Phonemic analysis and synthesis of pseudowords]. Unpublished manuscript.
Verhagen, W., Aarnoutse, C., & Van Leeuwe, J. (2006). *Predictoren voor beginnende geletterdheid* [Predictors of early literacy]. Pedagogische Studiën, 83, 226-245.
Verhagen, W., Aarnoutse, C., & Van Leeuwe, J. (2008). Phonological awareness and naming speed in the prediction of Dutch children's word recognition. *Scientific Studies of Reading, 12,* 301-324. doi:10.1080/10888430802132030
Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading, 4,* 313-330. doi:10.1207/S1532799XSSR0404_4
Verhoeven, L., & Aarnoutse, C. (1999). *Tussendoelen beginnende geletterdheid. Een leerlijn van groep 1 tot en met 3* [Intermediate goals beginning reading. A learning route from kindergarten to grade 1]. Nijmegen, The Netherlands: Expertisecentrum Nederlands.
Verhoeven, L., & Van Leeuwe, J. (2003). Ontwikkeling van decodeervaardigheid in het basisonderwijs [Development of decoding skill in primary education]. *Pedagogische Studiën, 80,* 257–271.
Vernooy, K. (2007). Leesonderwijs nader bekeken. Technisch lezen, woordenschat en leesstrategieën [A closer look at reading education. Technical reading, vocabulary and reading strategies]. Utrecht, The Netherlands: PO-Raad .
Venezky, R. (1978). Two approaches to reading assessment: A comparison of apples and oranges. In S. Pflaum-Connor (Ed.), *Aspects of reading instruction.* Berkeley, Cal.: McCutcheon.
Vygotsky, L. S. (1978). *Mind in society.* Cambridge, MA: MIT-Press.
Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101,* 192-212. doi:10.1037//0033-2909.101.2.192
Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & Rashotte, C. A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology, 85,* 83-103. doi:10.1037//0022-0663.85.1.83
Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). The development of reading related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology, 30,* 73-87. doi:10.1037/0012-1649.30.1.73
Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33,* 468 – 479. doi:10.1037/ 0012-1649.33.3.468
Webb, N. (1992). Testing a theoretical model of student interaction and learning in small groups. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in co-operative groups* (pp. 102 – 199). Cambridge, UK: Cambridge University Press.
Weber, G. (1971). *Inner-city children can be taught to read.* Washington, DC: Council for Basic Education.
Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education.* Cambridge, UK: Cambridge University Press.
Wentink, H., & Verhoeven, L. (2001). *Protocol leesproblemen en dyslexie* [Protocol reading problems and dyslexia]. Nijmegen, The Netherlands: Expertisecentrum Nederlands.
Wentink, H., & Verhoeven, L. (2004). *Protocol leesproblemen en dyslexie voor groep 5-8.* [Protocol reading problems and dyslexia for grade 3-6]. Nijmegen, The Netherlands: Expertisecentrum Nederlands.
Williamson, G. L., Applebaum, M., & Enpanchin, A. (1991). Longitudinal analyses of academic achievement. *Journal of Educational Measurement, 28,* 61-76. doi:10.1111/ j.1745-3984.1991.tb00344.x
Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes, and reading: A longitudinal study in average and impaired readers. *Child Development, 57,* 988-1000. doi:10.2307/1130373
Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91,* 415-438. doi:10.1037/0022-0663. 91.3.415
Wentink, H., & Wouters, E. (2005). *Protocol leesproblemen en dyslexie voor het speciaal basisonder wijs* [Protocol reading problems and dyslexia for special education]. Nijmegen, The Netherlands: Expertisecentrum Nederlands.

Samenvatting (Dutch summary)

Samenvatting (Dutch summary)

Lezen is een essentiële vaardigheid om goed te kunnen functioneren in onze maatschappij. Eén van de belangrijkste taken van het basisonderwijs is dan ook om alle kinderen te leren lezen. Kinderen leren lezen is een complexe taak. Het doel van het aanvankelijk leesonderwijs is om leerlingen te helpen zich de vele uitdagingen van het geschreven woord eigen te maken, zoals de kennis van het alfabetische systeem, het decoderen van woorden en het construeren, integreren en onthouden van de betekenis van woorden in een tekst.

Deze dissertatie richt zich op effectief leesonderwijs, meer specifiek op de ontwikkeling van beginnende geletterdheid. Beginnende geletterdheid is, na de fase van ontluikende geletterdheid en voorafgaand aan de fase van gevorderde geletterdheid, de leesontwikkelingsfase die loopt van het begin van groep 1 tot het einde van groep 3. Tijdens de fase van beginnende geletterdheid staan twee doelen centraal, namelijk het ontdekken van het alfabetische principe en het automatiseren van woordherkenning (Verhoeven & Aarnoutse, 1999). Woordherkenning is het proces van het omzetten van een reeks letters in klanken voor het identificeren van een woord (Coltheart, 1978).

De ontwikkeling van woordherkenning is zowel nationaal als internationaal veelvuldig onderzocht waarbij fonemisch bewustzijn, letterkennis en benoemsnelheid van onder andere letters en cijfers als belangrijke leesvoorwaarden worden genoemd (Aarnoutse, van Leeuwe, Voeten, & Oud, 2001; Bowers & Swanson, 1991; Hansen & Bowey, 1994; Näslund & Schneider, 1996; National Early Literacy Panel, 2008; Verhoeven & van Leeuwe, 2003; Wagner & Torgesen, 1987). Deze onderzoeken geven echter onvoldoende inzicht in welke instructiebenadering het beste de ontwikkeling van woordherkenning stimuleert. Het blijkt dat verschillende instructiebenaderingen verschillende effecten laten zien. Eén van de belangrijkste vragen die de verschillende instructietheorieën oproepen is of kennis dient te worden 'aangeboden' of te worden 'gegenereerd' (Rosenshine, Meister, & Chapman, 1996). De vraag van kennis aanbieden of kennis genereren is het vertrekpunt geweest van deze dissertatie. Om dit dilemma van 'aanbieden' versus 'genereren' te overstijgen is een derde weg gevonden die leidde tot de instructiebenadering Guided Co-Construction (Begeleide Co-Constructie). (Hardman, 2008; Mercer, 1995; Terwel, van Oers, van Dijk, & van den Eeden, 2009; van Schaik, Terwel, & van Oers, 2014). Guided Co-Construction (GCC) bevat de volgende drie elementen: 'Guided' verwijst naar de begeleidende en ondersteunende rol van de leerkracht, 'Co' refereert aan het samen leren en 'Construction' betreft de ontwikkeling van nieuwe inzichten en vaardigheden waarbij reeds verworven kennis en ervaringen worden benut.

In deze dissertatie is het effect van twee instructiebenaderingen op de ontwikkeling van woordherkenning onderzocht: een aanbiedende benadering Direct Instruction (Directe Instructie) en een genererende benadering met een coöperatieve leercomponent GCC. Direct Instruction (DI) is een instructiebenadering dat gebruikt maakt van gestructureerde lesfasen. Waarbij de activiteiten van de leerlingen en de leerkracht zorgvuldig omschreven zijn. Reeds verricht onderzoek toont aan dat DI een effectieve manier is om kinderen te leren lezen, met name de kinderen die het risico lopen leesproblemen te ontwikkelen (Adams, 1990; Bus & van IJzendoorn, 1999; Chall, 1996; Ehri, Nunes, Stahl, & Willows, 2001; Hattie, 2008). Recente reviews en meta-analyses tonen aan dat ook genererende instructiebenaderingen met een coöperatieve leercomponent bijzonder effectief zijn om kinderen te leren lezen (Raudenbush, 2009; Slavin, Lake, Chambers, Cheung, & Davis, 2009).

De overkoepelende vraag die in deze dissertatie wordt beantwoord is: Wat zijn de effecten van de twee instructiebenaderingen: Direct Instruction en Guided Co-Construction op de ontwikkeling in woordherkenning van groep 3 leerlingen en door welke factoren wordt de ontwikkeling van woordherkenning van leerlingen beïnvloed? De onderzoeksvragen die hieruit voortvloeien zijn:

1. Welke instructiebenadering (DI of GCC) is het meest effectief om de ontwikkeling van woordherkenning in groep 3 te stimuleren?
2. Welke leesvoorwaarden zijn belangrijk voor de ontwikkeling van woordherkenning?
3. Hoe ontwikkelt woordherkenning zich bij leerlingen in groep 3?

Om deze drie vragen te beantwoorden is de ontwikkeling van woordherkenning van 178 leerlingen uit groep 3 in een longitudinale studie onderzocht.

Guided Co-Construction versus Direct Instruction

In hoofdstuk 2 is onderzocht welke instructiebenadering het beste de ontwikkeling van woordherkenning stimuleert. Er is een veldexperiment uitgevoerd met een pre-test/post-test controlegroep-design. Twee varianten van het leesprogramma 'Veilig Leren Lezen' werden in twee verschillende groepen geïmplementeerd: de DI in de controlegroep en de GCC in de experimentele groep. In beide condities werden de leerkrachten getraind in de uitvoering van de verschillende instructiebenaderingen en materialen ontwikkeld en aangeboden om zowel DI als GCC te faciliteren.

Om te bepalen welk instructiebenadering het meest effectief was voor de ontwikkeling van woordherkenning, is de ontwikkeling van woordherkenning van 178 kinderen uit groep 3 met elkaar vergeleken. De meer specifieke vraag was: Is het beter om leerlingen tijdens het aanvankelijk leesonderwijs letter-klank relaties in van tevoren vastgelegde woorden aan te bieden (DI), of is het juist beter leerlingen te ondersteunen en te helpen zelf hun eigen letter-klank relaties en woorden te analyseren en te maken in samenwerking met medeleerlingen onder de begeleiding van de leerkracht (GCC)? Bovendien is onderzocht of leerlingen die verschillen qua sociaal-culturele achtergrond verschillend profiteren van de twee instructiebenaderingen.

De hypothese voor dit onderzoek was dat de woordherkenningsvaardigheden van de leerlingen in de GCC-groep beter zouden zijn dan die van de leerlingen in de DI-groep (Terwel, van Oers, van Dijk, & van den Eeden 2009; van Dijk, van Oers, & Terwel, 2003). Tevens werd verwacht dat leerlingen met een Nederlandse achtergrond meer van GCC zouden profiteren en dat leerlingen uit families met een buitenlandse achtergrond meer baat zouden hebben van DI (Leseman & De Jong, 1998).

Uit de eind-analyse met 'herhaalde metingen' blijkt dat de leerlingen in de GCC-groep over het geheel gezien beter presteren dan de leerlingen in de DI-groep. Echter, uit deze analyse komt ook naar voren dat de verschillen tussen beide groepen in de loop van het jaar afnemen en dat de DI-leerlingen aan het eind van het jaar hun tegenvoeters in de GCC-conditie vrijwel inhalen. Bovendien is er een differentieel effect gelet op de sociaal-culturele achtergrond. Leerlingen met een Nederlandse achtergrond profiteren meer van GCC terwijl leerlingen uit families met een buitenlandse achtergrond beter presteren in de DI-conditie. De grootste vooruitgang boeken de DI-leerlingen met een buitenlandse achtergrond. Als we afzonderlijk kijken naar de 109 Nederlandse leerlingen, dan blijven de prestatieverschillen tussen de condities nagenoeg constant gedurende het gehele jaar.

Implementatie

In hoofdstuk 3 is de kwaliteit van de implementatie van de twee instructiebenaderingen (DI en GCC) geanalyseerd. De onderzoeksvraag was of de twee benaderingen geïmplementeerd zijn zoals bedoeld. Dat wil zeggen, lieten de activiteiten van de leerlingen en de leerkrachten in de DI-groep meer DI-kenmerken zien dan de activiteiten van de leerlingen en de leerkrachten in de GCC-groep? En omgekeerd, lieten de leerlingen en de leerkrachten in de GCC-groep meer GCC-kenmerken zien dan de activiteiten van de leerlingen en de leerkrachten in de DI-groep? Om deze vragen te beantwoorden is het implementatieproces van de twee instructiebenaderingen zorgvuldig beschreven. Tevens zijn op systematische wijze observatiegegevens in beide groepen via de time-sampling methode verzameld en is de kwaliteit van de implementatie van beide groepen vergeleken. De resultaten toonden aan dat zowel de activiteiten van de leerkrachten als die van de leerlingen van elkaar verschilden in de beoogde richting.

Voorspellers van woordherkenning

Hoofdstuk 4 richtte zich op de vraag hoe goed woordherkenning van kinderen kan worden voorspeld. Uit onderzoek blijkt dat verschillende cognitieve factoren bijdragen aan het leren lezen (Adams, 1990; Verhagen, 2009) zoals het fonemisch bewustzijn (Aarnoutse, 2004), letterkennis als voorwaarde om het alfabetische principe te begrijpen (Bowey, 2005) en benoemsnelheid (Verhagen, 2009). In deze studie is het effect van fonemisch bewustzijn, letterkennis en benoemsnelheid op de woordherkenning na zes maanden leesonderwijs in groep 3 onderzocht.

Resultaten laten zien dat alle leesvoorwaarden een rol spelen in de snelheid van woordherkenning. De multiple regressie analyse waarin een set van achtergrondvariabelen is opgenomen toont dat het model 53% verklaart van de variantie van de snelheid van woordherkenning. Echter alleen letterkennis en de benoemsnelheid van cijfers hadden een significant direct effect. De conclusie is dan ook dat letterkennis en benoemsnelheid van cijfers het beste gebruikt kunnen worden om te voorspellen welke kinderen er mogelijk achter zullen blijven in hun leesvaardigheid.

Ontwikkeling van woordherkenning

In hoofdstuk 5 is de ontwikkeling van woordherkenning van leerlingen uit groep 3 onderzocht. Volgens verschillende onderzoekers laten kinderen met verschillende capaciteiten verschillende leesresultaten zien (Stanovich, 1986; Williamson, Appelbaum, & Enpanchin, 1991). In hoofdstuk 5 is onderzocht hoe leerlingen van groep 3 zich ontwikkelen in woordherkenning, hoe deze ontwikkeling verloopt voor zwakke en niet-zwakke lezers en of leerlingen vastzitten aan de classificatie 'zwakke en niet-zwakke' lezers?

Resultaten tonen aan dat de ontwikkeling van woordherkenning van de leerlingen in groep 3 bijna lineair is en dat zwakke lezers en niet-zwakke lezers zich parallel van elkaar ontwikkelen. Tevens laten de resultaten zien dat de leerlingen niet vastzitten aan de classificatie die aan het begin van het jaar gemaakt is. Ongeveer 40% van de lezers die in november als zwak geclassificeerd waren, waren in mei als niet-zwakke lezers geclassificeerd. En ongeveer 10% van de lezers die in november als niet-zwakke lezers geclassificeerd waren, waren in mei zwakke lezers.

Aanbevelingen voor toekomstig onderzoek

In hoofdstuk 2 is woordherkenning gekozen als belangrijkste variabele om de effecten van twee instructiebenaderingen (DI en GCC) te onderzoeken. Het is zinvol in vervolgonderzoek ook naar andere variabelen te kijken, zoals bijvoorbeeld woordenschat, begrijpend lezen en leesmotivatie. Ook kan worden gekeken naar coöperatieve vaardigheden zoals het luisteren naar anderen en het perspectief van anderen proberen te begrijpen (Ivey, 1994).

Van woordherkenning kan zowel de nauwkeurigheid als de snelheid gemeten worden. Omdat de snelheid van woordherkenning vaak het probleem is van kinderen die moeite hebben met het leren lezen, is er in de studies van deze dissertatie voor gekozen om de snelheid te meten. In een toekomstig onderzoek zou ook de nauwkeurigheid van woordherkenning onderzocht kunnen worden.

Tevens is er breder onderzoek nodig om meer inzicht te krijgen op het stimuleren van de leesontwikkeling van leerlingen en op het voorkomen van uitval bij het lezen. Toekomstig onderzoek zou zich daarom kunnen richten op: het voorkomen van leesproblemen, leesmotivatie, effectieve leesmethoden en effectieve leerkrachtvaardigheden.

Implicaties voor de onderwijspraktijk

De resultaten van deze dissertatie leveren enkele implicaties voor de praktijk op, die helpen de ontwikkeling van woordherkenning te stimuleren en uitval te voorkomen. Op basis van de resultaten van de eerste twee studies blijkt dat zowel DI als GCC binnen het aanvankelijk leesonderwijs in een gewone klassencontext uitvoerbaar zijn. GCC is over het geheel gezien het meest effectief om de ontwikkeling van woordherkenning te stimuleren. Echter, in de loop van het jaar worden de verschillen in gemiddelde prestaties tussen leerlingen in beide instructiegroepen steeds kleiner. Gedurende deze periode is wel extra aandacht nodig voor leerlingen afkomstig uit families met een buitenlandse achtergrond omdat zij meer profiteren van DI dan van GCC.

Daarnaast is het van belang om aan het eind van groep 2 letterkennis en benoemsnelheid van cijfers, te toetsen. Dit kan vroegtijdig problemen in de ontwikkeling van woordherkenning opsporen. Deze signalering maakt het mogelijk in een vroeg stadium interventies te plegen om verdere stagnatie te voorkomen.

Tevens is het van belang om in groep 3 dagelijks ruim voldoende tijd te besteden aan het aanvankelijk leesonderwijs met extra tijd voor de zwakke lezers.

Tenslotte heeft de laatste studie aangetoond dat leerlingen die in eerste instantie gecategoriseerd zijn als 'zwakke lezers' niet altijd 'zwakke lezers' blijven en omgekeerd. Dit betekent dat als de leerkracht van groep 3 zich vasthoudt aan de categorisering die aan het begin van het schooljaar gemaakt is, sommige leerlingen niet de begeleiding krijgen die ze nodig hebben. Daarom is het belangrijk om de woordherkenningsvaardigheden regelmatig te toetsen om naar aanleiding van die resultaten passende acties te ondernemen. Regelmatige toetsing is echter niet alleen in groep 3 van belang, maar ook in de daarop volgende groepen.

Curriculum Vitae

Mirjam Snel was born in Zaandam on 18 July 1968. Following secondary school, she studied primary education in the Hague at the Haagse Katholieke Pabo. She completed her degree in 1991 and went on to teach first grade in Voorhout at the Antoniousschool where she was served as an internal supervisor and (adjunct) director. In 1994, she started her Education study at the University of Leiden and completed this in 1997. Thereafter, she started her work at the Hogeschool Domstad (which became Utrecht University of Applied Sciences). She took on numerous coordinating and development task in addition to her work as a primary education teacher trainer there. In addition, in 2006, Mirjam started on a doctoral research trajectory, which she completed with her dissertation on the development of beginning reading in 2014. Since 2013, she has been Assistant Professor of Language at the Utrecht University of Applied Sciences. In addition to her coordinating and teaching tasks there, she cooperates with primary school teachers and primary education schools on the conduct of research into feedback and instructional skills in order to improve the curriculum for primary education training.