Reading for meaning

*The effects of Developmental Education on motivation and achievement in reading informative texts in primary school*

van Rijk, Y.; de Mey, L.; de Haan, D.; van Oers, B.; Volman, M.

*Published in:*
Research Papers in Education

*DOI:*
10.1080/02671522.2016.1225789

*Link to publication*

Creative Commons License (see https://creativecommons.org/use-remix/cc-licenses):
CC BY-NC-ND

*Citation for published version (APA):*
Reading for meaning: the effects of Developmental Education on motivation and achievement in reading informative texts in primary school

Yvonne van Rijk, Langha de Mey, Dorian de Haan, Bert van Oers & Monique Volman

To cite this article: Yvonne van Rijk, Langha de Mey, Dorian de Haan, Bert van Oers & Monique Volman (2017) Reading for meaning: the effects of Developmental Education on motivation and achievement in reading informative texts in primary school, Research Papers in Education, 32:3, 333-352, DOI: 10.1080/02671522.2016.1225789

To link to this article: http://dx.doi.org/10.1080/02671522.2016.1225789

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 02 Sep 2016.

Article views: 177

View related articles

View Crossmark data
Reading for meaning: the effects of Developmental Education on motivation and achievement in reading informative texts in primary school

Yvonne van Rijk, Langha de Mey, Dorian de Haan, Bert van Oers and Monique Volman

Abstract
Content-oriented reading interventions that focus on the integration of motivational enhancement and strategy instruction have been found to have positive effects. Developmental education (DE) in the Netherlands is an innovative content-oriented approach in which reading is an integral part of an inquiry-oriented curriculum. Reading for meaning is a central pursuit, and strategy instruction is functionally integrated. This study differs from previous studies in three aspects. Firstly, instead of interventions, two types of existing practices were compared: the DE approach and a textbook-driven programmatic instruction approach (PI). Secondly, this study accounted for classroom influences by conducting multilevel analyses. Thirdly, control variables, i.e. ethnic background, home language, SES, non-verbal IQ, gender, vocabulary and decoding skills, were taken into account. The effects of both approaches were investigated in terms of reading comprehension, knowledge of reading strategies and reading motivation. In a pretest–posttest natural two-group design, tests and questionnaires were administered to 570 grade 4 students in 24 schools. The outcomes that resulted from the DE approach were as good as those from the PI approach. These results are discussed in relation to previous studies that have reported better outcomes of content-oriented reading approaches than traditional and strategy reading approaches.

Introduction
Reading for meaning has been revealed to be essential to the effectiveness of teaching reading comprehension (Snow 2002). It is the point of departure for content approaches of reading. The central pursuit of a recent innovative content approach, Concept-Oriented Reading Instruction (CORI), which was developed by Guthrie and Wigfield (2000), is to foster an in-depth understanding of the conceptual content of the text through the avenue of active engagement of the student. Extensive research has revealed that CORI interventions have positive effects on reading motivation and reading comprehension compared to approaches that are based on traditional or mere strategy instruction (Guthrie and Klauda 2014; Guthrie, McRae, and Klauda 2007; Guthrie et al. 2004; Wigfield et al. 2008).

Contact Yvonne van Rijk y.van.rijk@hva.nl

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-ND/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.
Content approaches of reading may differ in the scope of their curriculum. A content approach like developmental education (DE) in the Netherlands has a broad scope, being embedded in a pedagogical concept which covers all learning activities. Teaching reading in DE schools (about 5% of the primary schools in the Netherlands), is integrated in a play and inquiry curriculum from Kindergarten to the highest grade of primary school, and related to other learning activities within topical themes of interest. Since reading activities in DE are inspired by the CORI characteristics (see below), there is reason to assume that DE will also yield positive results compared to the regular textbook-driven programmatic approach in Dutch schools in which strategy instruction has a main role (see below). However, other Dutch studies that were inspired by CORI have failed to produce convincing positive effects. In Aarnoutse and Schellings’ (2003) intervention study, in which students in grade 3 were challenged to solve a self-formulated problem by reading all types of written resources, the experimental group significantly outperformed the control group in reading motivation and knowledge of reading strategies. However, no effect was found on reading comprehension. Droop et al. (2012, 2016) researched an intervention in grades 3 and 4, in which strategy instruction was integrated into a meaningful context. Compared to current Dutch programmes for reading comprehension, a positive effect was found on knowledge of reading strategies, but no differences were found in reading motivation or reading comprehension of informative texts. In this study, we are interested to find out whether the positive effects of the CORI interventions will be found in the context of an existing, all-encompassing practice like DE.

A more practical motivation for our research is the scepticism in Dutch educational policy about innovative practices such as DE that do not programme the learning process in a detailed manner through strict adherence to textbooks. Although schools in the Netherlands are free to decide how curricula are taught, the pressure on schools to use a reading programme that is based on a textbook was increased by a parliamentary investigation committee that critically evaluated a number of educational innovations (Dijsselbloem 2008). Similarly, a decline in the Netherlands' position in PISA rankings (OECD 2010) has enhanced the emphasis on the use of textbook-driven programmes to teach basic reading knowledge and skills. Tendencies towards greater performance accountability in the Netherlands have encouraged schools to adopt reading programmes that are based on textbooks, as these programmes are better suited to the standardised tests that schools are required to use.

Hence, the aim of our research was to establish the effects of reading for meaning as enacted in two existing approaches: DE and the textbook-driven programmes to teach reading, which we will refer to as PI.

**Theoretical background**

*Content-oriented and strategy approaches to reading comprehension*

Meaning construction is considered the core of reading. In the review of the Rand Reading Study Group, reading is considered ‘the process of simultaneously extracting and constructing meaning through interaction and involvement with text’ (Snow 2002, 11). In this review, a distinction is made between means for understanding, like the use of strategies, and the main purposes of reading: ‘gaining meaning and gaining knowledge’ (Snow 2002, 40). A challenge in reading instruction is to find the optimal balance between the instruction of the means, of which strategy instruction is most prominent, and the focus on content. The evidence is growing that both content and strategy approaches are effective in teaching reading comprehension. In the following, we shortly discuss the theoretical basis of these approaches, and then we go into the programmes investigated in this study. Regarding content approaches, we focus on CORI, since DE agrees to a great extent with this innovative content approach.

Content approaches to reading are based on theoretical models of text processing, emphasising the construction and integration of meaning in a coherent whole to gain knowledge, and on socio-cognitive perspectives involving shared meaning making (McKeown, Beck, and Blake 2009). In CORI, this theoretical model is included in an ‘engagement perspective’, in which engagement processes are
assumed to mediate the effect of instructional classroom practices on reading comprehension to a large extent. Reading engagement is defined as ‘interacting with text in ways that are both strategic and motivated’ (Guthrie, Wigfield, and You 2012, 602). The view is that, if reading practices in the classroom provide students with content that is meaningful to them and relevant to their lives, they are motivated to read. This motivation, conceived as student’s goals, values and beliefs with regard to the reading activity, is thought to energise and direct the cognitive behavioural engagement processes like the use of strategies involved in reading. Comprehension strategies are procedures and routines that readers themselves apply across a number of different texts (NICHHD 2000). The effortful and deliberate processes of strategic behaviour in their turn lead to understanding a text and, ultimately, to reading competence (see for a fine-grained discussion and evidence Guthrie, Wigfield, and You 2012). Motivation is enhanced by a number of motivational practices, such as (1) the integration of reading and strategy instruction in domain-specific content matters, (2) the provision of interesting texts to strengthen relevance and the value of reading, (3) text choices to support autonomy and intrinsic motivation, (4) positive feedback regarding comprehension for self-efficacy, (5) the organisation of collaboration to enhance social motivation in sharing cognitive activity (see for a fine-grained discussion and evidence Guthrie, Wigfield, and You 2012).

In the Netherlands content and strategy approaches are more or less recognisable in DE and PI methods, respectively. PI is characteristic of the current situation of the large majority of Dutch schools and involves both traditional elements (exercises for vocabulary and textual skills like pronominal procedures) and insights from theories of the strategy approach. The reading programmes in PI differ in type and number of reading strategies, time for reading, proportion of classroom instruction, individual exercises or collaboration and type of texts. However, Heesters et al. (2007) found no significant differences in effectiveness for these programmes. The common characteristic of all programmes is the prescribed, pre-planned character of the lessons relating to organisation, selected strategies, instruction and texts (see Figure 1). The reading programmes consist of a grade-levelled series of textbooks that are produced by an educational publisher. Teachers are provided with a teacher manual for guidance that largely prescribes how each lesson should be taught and students have booklets with texts and exercises. The pre-planned lessons are provided to the entire class and are scheduled on a weekly basis. PI is focused on direct instruction; however, the teacher’s manual mostly does not provide direct clues about the didactics of explanation and explicit modelling (see also Droop et al. 2016 for these practices in Dutch regular schools). It is common practice in all these programmes to instruct students to apply the reading strategies before, during and after reading. Motivating students is not a prominent issue since texts are fixed for all pupils. In order to make texts interesting, subjects are chosen which are supposed to fit in with the students’ everyday world.

In DE, reading is an integral part of an all-encompassing pedagogical approach in which agency and citizenship are the main objectives of education. The desire for meaning is considered to be the engine of all learning. DE is based on Cultural Historical Activity Theory as conceived by Vygotsky, Leont’ev and their followers (Lee and Smagorinski 2000; Van Oers et al. 2008; Wells and Claxton 2002). The context for development is the interpersonal activity, i.e. the students’ interactions in the context of social activities in which students bring in their own voices and histories. Development is seen as a gradual process of identity formation through participation in cultural practices (like for instance literacy practices) with more knowledgeable others.

DE seeks to facilitate the development of students’ identities as active and responsible participants in society (Van Oers 2009; Van Rijk et al., forthcoming; Pompert 2004; Wardekker 2012). The DE classroom is a community of learners in which inquiry is the core activity in the higher grades (cf.
Wells 2000). Reading informative texts is embedded in this inquiry curriculum as a resource for problem-solving; there are no reading activities that do not have a question or a purpose. Reading is seen as a form of communication that is driven by students’ research questions, and texts function as resources to find answers to those questions. Oral interaction and writing are inseparably connected to reading, and from the beginning in kindergarten, literacy makes sense to students (Pompert 2012). In DE, teachers are educational designers. They create meaningful thematic units in which all activities relate to the theme and emerging problems are interesting to students. Teachers seek to involve all students in reading texts that correspond to their self-generated problems, interests and reading levels. Students are encouraged to bring relevant texts from home or the internet into the classroom and to share their ideas and thoughts. They perform inquiries in collaborative expert groups and visualise their questions in ‘wall posters.’ They may interview experts in their own environments or invite them to the classroom, or may set up an experiment to find answers. Direct strategy instruction and teaching

<table>
<thead>
<tr>
<th>Programmatic Instruction</th>
<th>Developmental Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Reading is taught as a separate discipline in the curriculum. In addition to the reading programs, programs for geography, history and science are used, but these do not contain explicit reading instructions.</td>
<td>Reading is integrated in meaningful (thematic) content and activities in which historical, geographical and science perspectives are integrated.</td>
</tr>
<tr>
<td>2 Strategy instruction is a main focus of reading programs (knowledge and application), and is considered as a main mean for reading comprehension.</td>
<td>Content is the focus of reading, strategies are considered as functional tools for text comprehension.</td>
</tr>
<tr>
<td>3 Strategy instruction is programmed: strategies are taught following the order of the textbook</td>
<td>Strategy instruction is provided flexibly, when needed for students’ comprehension of a text.</td>
</tr>
<tr>
<td>Instruction is provided to the whole-class, following standard procedure: the teacher provides instruction, students answer written questions about the text. The teachers helps students individually.</td>
<td>The teacher provides instruction flexibly to the whole-class when the text’s content is relevant for all students, to small (expert) groups when strategies are needed for comprehension of the text’s content.</td>
</tr>
<tr>
<td>Instruction involves explicit instruction of strategies (known as ‘direct instruction’ in the Netherlands).</td>
<td>Instruction involves demonstration, modeling, and dialogue about solving problems regarding the content of the text.</td>
</tr>
<tr>
<td>4 Texts are fixed (textbook-driven), all students read the same texts</td>
<td>Texts are provided by the teacher or chosen by students. Text choice is based on interest (content), as resources to find answers to students’ questions.</td>
</tr>
<tr>
<td>5 Self-efficacy is enhanced by pre-programmed differentiation; positive feedback</td>
<td>Self-efficacy is enhanced by selection of appropriate texts by the teacher and positive feedback.</td>
</tr>
<tr>
<td>6 Students mostly work individually and sometimes collaboratively</td>
<td>Students often collaborate in discussions and presentation of findings of the expert groups.</td>
</tr>
<tr>
<td>7 Time is scheduled: 2 lessons of 45 minutes per week (N.B. both informative and fictional texts)</td>
<td>Reading is organized on a daily basis, integrated in other activities, without a fixed time schedule.</td>
</tr>
</tbody>
</table>

Figure 1. Characteristics of Pi and DE.
of vocabulary and textual skills is an integrated part of the activities, and whole classroom instruction is alternated with flexible instruction adjusted to the needs of the individual student or expert group. Students write down answers to their questions to inform each other. Classroom discussions are occasions for collaborative thinking about problems and the student’s answers, and inspire the students to revise their texts. Together with other products from the thematic unit, the revised texts are prepared for presentation, often to an audience of parents and students from other groups (Van Rijk et al., forthcoming). From this overview, it may be clear that DE incorporates motivational practices that are similar to those of CORI (Guthrie et al. 2004). In Figure 1, we describe the differences between PI and DE.

This research

This research aims to establish the effects of reading for meaning as enacted in two existing approaches, DE and PI. The focus is on informative texts, since this type of text is essential for students’ future school careers. On theoretical grounds, the overall assumption is that DE, with its focus on meaning, adds value to reading instruction and motivates students to read informative texts. On empirical grounds (CORI), the assumption is that DE, by providing a combination of motivation enhancement and strategy instruction, leads to better reading comprehension than PI, where motivation is not a prominent issue.

This paper aims to contribute to the insights from former studies in three ways. First, rather than examining an intervention, two types of existing practices were compared: DE characterised by teaching reading in which meaning and engagement is considered the core of reading, and PI by textbook-driven, prescribed and pre-planned instruction of strategies, text structures and vocabulary. Second, this study accounts for contextual (classroom) influences via multilevel statistical analyses. Third, control variables that may account for variance in students’ reading comprehension, motivation and knowledge of reading strategies were taken into account. Research has shown that IQ, gender and social and ethnic background are related to students’ reading motivation and/or reading comprehension in the upper grades of primary education. Van Elsäcker (2002) reported an effect of non-verbal IQ on reading comprehension. Recent studies have found that girls have more favourable intrinsic/autonomous motivations for reading than boys, and that students of ethnic minority groups have higher reading motivation (De Naeghel and Van Keer 2013; Hornstra et al. 2013). In contrast, ethnic minority backgrounds and home (native) languages may lead to smaller vocabularies among children for whom the language of the reading texts is a second language (De Naeghel and Van Keer 2013) which has an effect on reading comprehension (Babayiğit 2014; Snow 2002; Van Elsäcker 2002). Furthermore, decoding skill is generally known to influence reading comprehension (Snow 2002).

The main question of the present study was the following: With respect to informative texts, what are the effects of the DE approach compared to those of the PI approach in terms of reading comprehension, knowledge of reading strategies and reading motivation in fourth-grade primary education?

The research questions were as follows:

1. Do reading comprehension outcomes of students in the DE approach and students in the PI approach differ?
2. Does the knowledge of reading strategies among students in the DE approach and students in the PI approach differ?
3. Does the motivation for reading informative texts differ among students in the DE approach and in the PI approach?
4. How do variables that are related to reading (vocabulary and decoding skills) and to students’ background (non-verbal IQ, ethnic background–home language, SES–parental education level and gender) influence the outcomes for reading comprehension, strategy knowledge and motivation?
Methods

Design

This study employed a pretest–posttest natural two group design, because we wanted to find out how student reading comprehension and reading motivation would evolve over time, in the course of a school year, for two different approaches, i.e. DE and PI. The innovative DE approach is compared with the PI approach, which is default in the Netherlands.

The pretests were administered in the beginning of the school year (September and October), and the posttests were conducted at the end (May and June). At both measurement time points, tests of reading comprehension, decoding skills and reading vocabulary were administered to the students, and students completed questionnaires about reading motivation and knowledge of reading strategies. Non-verbal IQ tests were administered only at the pretest. No more than two sessions per day were scheduled to minimise test weariness.

Participants

As schools generally begin teaching students to read informative texts at the age of 9 or 10 in the Netherlands, this study was performed in grade 4. Fourth-grade students ($N = 570$) and their teachers from 24 schools participated (one class per school). The DE group ($n = 258$ students) consisted of 12 classes and had 16 teachers. Five classes combined students from the third and fourth grades, but only fourth graders participated in this study. The PI group ($n = 312$ students) consisted of 12 fourth-grade classes with 15 teachers. Schools and teachers voluntarily participated in this study, and informed consent was obtained from the students’ parents.

First, the schools for the DE sample were selected with the help of the register published by the Inspectorate of Education and acknowledged as DE-schools by experts in this field. The selection criterion was that schools had adopted the concept of DE. These schools have organised education into thematic units and provided meaningful text reading (see Figure 1 for the main characteristics of DE). The participating teachers were required to have at least two years of experience with DE in grades 3 and higher. We purposefully varied our sample population in terms of urban, suburban and rural situations, high and low SES areas and ethnic minority compositions.

Next, the PI sample was selected. The selection criteria were (1) the use of a modern reading programme that was based on textbooks that are currently used in Dutch primary education and (2) that the teachers had at least 2 years of experience in grades 3 and above.

The selection criteria of PI and DE were discussed with the principals and teachers during the recruitment process. Only schools that met the criteria were selected for participation.

The PI sample was matched to the DE sample at classroom level in terms of the following set of relevant characteristics: proportion of boys and girls, SES of students (parental education level), ratio of students from ethnic minority groups (parental country of origin) and percentage of another home language (the language predominantly spoken at home) than Dutch.

Data regarding the background characteristics of the students were collected from the school administration and included student gender, SES (parental education level) and ethnicity (parental country of origin) (Table 1). The SES categories corresponded to the following levels of the International Standard Classification of Education (UNESCO 2006): primary education, lower secondary education and upper secondary education. The Dutch Government provides school funding for students with less educated parents. We followed the government classifications, to distinguish the high SES group from other (i.e. middle and low) SES groups. To determine ethnicity, we asked students at the posttest to provide the country of origin of their parents because the schools used different systems for the registration of ethnicity. Since the extent to which the language spoken at home may be a predictor of reading comprehension, we also asked them about the language(s) spoken predominantly at home (Dutch, another language or a mix of both).
As we wanted to verify that reading instruction was provided as intended in both DE and PI, a questionnaire was used to determine the extent to which the participating teachers showed features typical of DE in their reading instruction of informative text. All teachers, of both DE and PI, have filled in the questionnaire. Additionally, during data collection, a researcher was present five or six times in every classroom. Informal observation of the classrooms and informal conversation with the teachers offered an extra opportunity to verify the distinctive features of specific reading programmes: the types of text (textbooks for PI, variety of meaningful texts for DE), the way in which teachers provided reading instruction (teacher manuals for PI, flexible instruction adjusted to students’ level from teachers for DE) and the classroom context (no extra texts visible in PI classrooms versus a variety of texts, questions (wall posters) and collaboration in DE classrooms). DE-teachers designed their teaching and PI-teachers adhered closely to the programme, in accordance with the selection criteria.

**Instruments**

All tests used in this study were national or international standardised tests for which the validities were previously established.

*Reading motivation* was measured with a questionnaire that consisted of 46 items that were spread over seven subscales that were validated or tested in previous research. Students were asked to indicate

<table>
<thead>
<tr>
<th>Table 2. Subscales of the reading motivation questionnaire.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subscale</strong></td>
</tr>
<tr>
<td>Curiosity</td>
</tr>
<tr>
<td>Preference for challenge</td>
</tr>
<tr>
<td>Extrinsic motivation</td>
</tr>
<tr>
<td>Intrinsic value</td>
</tr>
<tr>
<td>Utility value</td>
</tr>
<tr>
<td>Self-efficacy</td>
</tr>
<tr>
<td>Instrumental motivation</td>
</tr>
</tbody>
</table>
the extent to which each item applied to them on a five-point scale (1 = does not suit me at all; 5 = suits me very well), or a three-point scale (1 = always, 2 = sometimes, 3 = never), depending on the origin of the scale.

The subscales used are displayed in Table 2:

The curiosity, preference for challenge and extrinsic motivation subscales originated from the Motivations for Reading Questionnaire (MRQ) (Guthrie 2010). We used the translation for the Dutch context (Förrer and Van de Mortel 2010). The other subscales have been developed and tested by Dutch researchers: intrinsic value, utility value, self-efficacy by De Milliano (2013) and instrumental motivation by Tellegen and Frankhuisen (2002).

For the present sample, reliability of the MQR was poor (0.64) for preference for challenge and extrinsic motivation. These two subscales were not retained in further analyses. For the five subscales retained reliability was estimated between 0.68 and 0.78, which is acceptable (Field 2009; Sijtsma 2009).

Reading comprehension was assessed using the Progress in International Reading Literacy Study (PIRLS) of 2006, for students aged 9–10 in 40 countries (Mullis et al. 2007). We used the part of the test that consists of a booklet with expository texts (three texts about Searching for food). Students answered 15 questions, 8 of which were multiple choice and 7 of which were open-ended questions. For three open-ended questions a score of 2 points could be obtained, making the overall maximum score 18. The questions assessed the following four categories: (1) the ability to focus on and retrieve explicitly stated information, (2) the creation of straightforward inferences, (3) the interpretation and integration of ideas and information and (4) the examination and evaluation of content, language and textual elements. An open-ended question was, for example: ‘What is similar in the way ants and pill bugs find their food?’ The open-ended questions were scored by two fellow researchers, using the PIRLS scoring guide. Answers were discussed to obtain consensus. For the present sample, reliability outcomes were moderate (Cronbach’s $\alpha$ was 0.68).

Knowledge of reading strategies was assessed via a questionnaire that was developed by Brand-Gruwel (1995). This questionnaire consisted of 22 multiple choice questions regarding strategies that are to be applied before, during and after reading, such as goal setting and predicting, question answering and summarising. Questions were of the type ‘What is best to do … (e.g. before reading)?’ Students had to choose the best option out of three: the most adequate strategy, a less adequate strategy and an irrelevant strategy. For the present sample, Cronbach’s $\alpha$ was 0.63.

Decoding skills were assessed using the three-minute test, which is part of the Cito pupil monitoring system used by nearly all Dutch schools (Verhoeven 1995). This standardised test measures decoding skills in terms of the rapid recognition and naming of unrelated words. The test is administered individually. It consists of a set of three cards with increasing degrees of difficulty. The first card has 150 monosyllabic words (e.g. hot), the second card has 150 more complex monosyllabic words (e.g. spring, worst) and the third card has 120 polysyllabic words. For every card, students are required to read as many words as possible in one minute. The test score consists of the total number of words that were correctly read across all three cards, with a maximum score of 470. For the present sample, Cronbach’s $\alpha$ was estimated at 0.93.

Reading vocabulary was measured using a standardised test (Verhoeven and Vermeer 1995). This test consists of 30 multiple choice items. From among four possible answers, students were required to choose the meaning of the word that was underlined in a sentence. For the present sample, Cronbach’s $\alpha$ was 0.78.

Non-verbal IQ was measured using the Raven SPM (Raven, Raven, and Court 2000). This test measures students’ reasoning abilities and consists of 60 items that are listed by order of difficulty. Each item presents a set of geometrical figures of which one figure was missing. The missing item must be selected from a set of six or eight answers. The test was administered to the whole group. For the present sample Cronbach’s $\alpha$ was 0.79.

The teacher questionnaire was used to verify that reading instruction was provided as intended in both DE and PI. The questionnaire consisted of 26 statements on a five-point Likert scale. The maximum sum score was 130. The statements were divided into six categories. The first category was about
the programmatic or non-programmatic character of the approach (e.g. ‘I use a pre-programmed textbook for reading comprehension’). This was followed by five sets of statements, each covering one of the motivational components that were distinguished by Guthrie et al. (2004) and have been incorporated into DE – integration of reading instruction in subject matter, providing interesting texts, choice of text, collaboration, fostering successful reading experiences by providing appropriate texts and positive feedback (self-efficacy). Items were for example ‘In my class students use information from texts read for other learning activities’ or ‘In my class students may choose informative texts to read’.

Data analyses

The student data in the present study had a two-level hierarchical structure; i.e. students (level 1, \(N = 570\)) were nested within classrooms (level 2, \(n = 24\)). As only one class per school participated, the classroom level was equal to the school level. The assumption of the independence of the observations was tested. We deduced from the intra-class correlations (ICCs) that multilevel modelling was the most suitable approach for data analysis (see Table 5). We used the mixed-model procedures of SPSS (20.0) with maximum likelihood estimation\(^1\) (Hox 1995).

As we were interested in effects of the educational approaches on the outcome variables of reading motivation, reading comprehension and knowledge of reading strategies, each of these variables was separately analysed (Table 6) using a step-up model building strategy (West, Welch, and Galecki 2007). Standardised scores (z-scores) were used for analyses and effect sizes were calculated using Cohen’s \(f^2\), which is a standardised measure for hierarchical linear modelling (Selya et al. 2012). According to Cohen’s (1988) guidelines, \(f^2 \geq 0.02\) represents small effect size, \(f^2 \geq 0.15\) represents medium effect size and \(f^2 \geq 0.35\) represents large effect size. For testing the results, a 95% confidence level was used. For purposes of further exploration of differential effects in DE and PI, analyses were also performed on the DE group and the PI group separately.

All analyses begin with intercept-only (or unconditional) models in which the intercepts (mean levels) of the dependent variables were estimated and the random effects (or estimations of variance) of the following two components were also estimated: classroom means (i.e. classroom intercepts) and the individual errors (i.e. residuals). Subsequently, the predictors (co-variates) at level 1 (student) were included one by one, while testing their fixed effects and relevant interactions. We investigated multi-collinearity; i.e. strong correlations between the predictors. No multi-collinearity was found based on the criterion that a correlation of 0.80 or higher may imply collinearity, and the calculated tolerance indexes (VIF) did not produce concern (Field 2009). As no multi-collineairties were found among any of the predictors, the order of the inclusion of the co-variates did not affect the estimation of the parameters. First, the pretest measure of the outcome variable was included as a co-variate to increase the power of the analyses (Lipsey and Hurley 2009). Next, the variables related to reading (that were analysed separately as outcome variables) were included in the model. Since every outcome variable may also be a predictor of another outcome variable, these variables are included in the regression, along with the possible confounders. The confounders at level 1 were vocabulary, decoding skills, non-verbal IQ, SES, ethnicity and gender. At level 2, first of all the approach (DE versus PI) was added to the model to test the effects of DE and PI. Confounding effects were expected from the pretest at classroom level, SES (i.e. the proportion of low and middle SES students in the classroom) and ethnicity (i.e. the proportion of non-native Dutch and mixed ethnicity students in the classroom). Each included co-variate was tested for its contribution to the model fit and its effect on the outcome variable. The calculated deviance statistics (−2 log likelihood) of the different models were compared and tested with a \(\chi^2\) difference test to draw conclusions about the improvement of the model. A positive conclusion led to the inclusion of the predictor in the model. The final models were constructed on the basis of the parsimony principle. More parsimonious models were created by removing non-significant predictors whose removal did not negatively affect the model fit.

Inspection of the missing values revealed that, for some co-variates, the missing data points were not randomly distributed and were also substantial (>5%). Multiple imputation was applied, as this
technique is favourable for parameter estimation (Tabachnick & Fidell, 2011). In addition to the original data-set, five imputed sets were created, using the multiple imputation procedure in SPSS 20. The outcomes of the ‘pooled’ imputed set were used in subsequent analyses.

Furthermore, to simplify the interpretation of the results, we centred the non-categorical predictors (i.e. reading comprehension, knowledge of reading strategies, reading motivation, decoding skills, vocabulary and non-verbal IQ). We also dichotomised the categorical variables SES (in high SES versus middle and low SES) and ethnicity (in native Dutch parents and Dutch as the home language versus non-native Dutch and mixed parents and home language). For reading motivation, which was originally a categorical variable, the sum of the average scale scores was used.

The results of the multilevel analyses are displayed in Table 6.

**Results**

This section begins with the descriptions of the means and standard deviations of all of the tests segregated by group and measurement occasion (Table 3).

**Preliminary analyses**

An ANOVA revealed that responses to the teacher questionnaire, by teachers in the DE and the PI approach, differed significantly in three aspects: the use of a programme that was based on a textbook, $F(1) = 10.801; p = 0.03$, the provision of choice of texts, $F(1) = 19.087; p = 0.02$ and the integration of reading instruction with content, $F(1) = 12.106; p = 0.02$. These results revealed that the teachers were teaching according to the approaches of their groups. No significant differences were found regarding teachers’ answers about three other aspects. These were providing interesting texts, $F(1) = 0.692; p = 0.415$ (DE: $M = 3.67$, SD = 0.75; PI: $M = 3.43$, SD = 0.54), organising collaboration, $F(1) = 0.822; p = 0.724$ (DE: $M = 4.13$, SD = 0.83; PI: $M = 4.33$, SD = 0.71) or fostering successful reading experiences by providing appropriate texts and positive feedback (self-efficacy), $F(1) = 0.360; p = 0.554$ (DE: $M = 3.56$, SD = 0.73; PI: $M = 3.26$, SD = 0.85). The contrasts of these approaches were considered sufficient for further analyses (see Table 4).

Next, the dependencies of the data were tested. Table 5 presents the results of the intercept-only models of the outcome variables; i.e. reading comprehension, knowledge of reading strategies and reading motivation. As previously discussed, the intercept-only models were used to test the dependencies of the data.

| Table 3. Means and standard deviations of the tests by group and measurement occasion. |
|-----------------------------------------------|-----------------------------------------------|
| Pretest                                      | Posttest                                      |
| **DE group**                                 | **PI group**                                  |
| $n$                                          | $M$   | $SD$ | $n$   | $M$   | $SD$ | $n$   | $M$   | $SD$ |
| Reading motivation<sup>a</sup>               | 202   | 12.47| (1.58)| 215   | 12.56| (1.44)| 235   | 12.44| (1.49)| 271   | 12.47| (1.38)|
| Reading comprehension<sup>b</sup>            | 248   | 7.33 | (3.09)| 243   | 9.26 | (3.43)| 293   | 6.93 | (2.96)| 291   | 9.62 | (3.41)|
| Knowledge of reading strategies<sup>c</sup>  | 253   | 35.99| (4.30)| 243   | 38.12| (3.73)| 292   | 34.71| (4.26)| 292   | 36.88| (4.04)|
| Decoding skills<sup>d</sup>                  | 252   | 225.63| (50.81)| 252   | 262.50| (48.81)| 293   | 229.06| (53.96)| 300   | 258.47| (49.38)|
| Vocabulary<sup>e</sup>                      | 249   | 17.52| (4.87)| 251   | 17.78| (5.25)| 296   | 18.39| (4.62)| 296   | 18.60| (4.97)|
| Non-verbal IQ<sup>f</sup>                   | 249   | 38.62| (8.30)| 300   | 39.93| (6.56)|

<sup>a</sup>Mean sum: range 0–19;  
<sup>b</sup>Total score: range 0–18;  
<sup>c</sup>Total score: range 0–44;  
<sup>d</sup>Total score: range 0–420;  
<sup>e</sup>Total score: range 0–30;  
<sup>f</sup>Total score: range 0–60.
Multilevel analyses

Table 6 presents the final multilevel models and reveals the effects of the approach (DE or PI) on the outcome variables of reading comprehension, knowledge of reading strategies and reading motivation. The final models presented in Table 6 reveal significant improvements over the intercept-only models presented in Table 5.

The effects on reading comprehension

The first research question was answered in the negative. The educational approach (DE or PI) did not significantly influence reading comprehension outcomes as seen in Table 6, which contains a blank for the approach.

As displayed in Table 6, reading motivation had a positive effect on reading comprehension ($b = 0.08, SE = 0.03, p < 0.05, f^2 = 0.02$).

Of the predictors, vocabulary ($b = 0.28, SE = 0.04, p < 0.01, f^2 = 0.09$) and non-verbal IQ ($b = 0.34, SE = 0.05, p < 0.01, f^2 = 0.06$), had a positive effect on reading comprehension. Decoding skills had no effect and did not improve the model fit, and were therefore not retained in the model.

The background variables did not have any effect. SES and ethnicity, however, remained in the model because they improved the model fit. We will explain this finding further on in this section.

Of the variables that were tested at classroom level (level 2) only the pretest had an effect on reading comprehension ($b = 0.12, SE = 0.05, p < 0.01, f^2 = 0.00$). An interaction effect was found between the educational approach and non-verbal IQ at student level ($b = −0.20, SE = 0.07, p < 0.01$). We explored the nature of this interaction by performing analyses on the DE group and the PI group separately. In both approaches (DE and PI), non-verbal IQ had an effect on reading comprehension, but in PI the effect size was considerably greater than in DE. This indicates that in PI students with high non-verbal IQ perform better on reading comprehension, whereas in DE, non-verbal IQ is less important for reading comprehension.

Further exploration showed some other differential effects (Appendix Tables A1 and A2). In PI, none of the reading variables had an effect on reading comprehension, whereas in DE, reading motivation
(b = 0.10, SE = 0.05, p < 0.05, \( f^2 = 0.01 \)) and knowledge of reading strategies (b = 0.14, SE = 0.05, p < 0.01, \( f^2 = 0.03 \)) had positive effects. This indicates that in DE, reading motivation and knowledge of reading strategies are more important for reading comprehension than in PI.

Of the predictors, vocabulary had a positive effect on reading comprehension in both approaches (DE: b = 0.21, SE = 0.06, p < 0.01, \( f^2 = 0.08 \)); PI: b = 0.31, SE = 0.06, p < 0.01, \( f^2 = 0.11 \)). In DE, decoding skills also had a positive effect on reading comprehension (b = 0.13, SE = 0.05, p < 0.05, \( f^2 = 0.05 \)). This indicates that in DE, students with good decoding skills perform well in reading comprehension.

Of the background variables, no effect was found in either approach. In DE, however, SES showed a medium effect size (b = −0.13, SE = 0.14, p > 0.05, \( f^2 = 0.28 \)), in contrast to PI. This indicates that low parental education level (SES) seems to have a negative effect on reading comprehension.
The effects on knowledge of reading strategies

The second research question was answered in the negative. As illustrated in Table 6, the educational approach (DE or PI) did not have an effect on the development of knowledge of reading strategies.

However, the analysis revealed that reading motivation exerted a positive effect on knowledge of reading strategies ($b = 0.12, SE = 0.04, p < 0.01, f^2 = 0.02$).

Among the other predictors, non-verbal IQ ($b = 0.13, SE = 0.05, p < 0.01, f^2 = 0.02$) and gender ($b = 0.23, SE = 0.08, p < 0.01, f^2 = 0.02$) affected knowledge of reading strategies. Students with a higher non-verbal IQ and girls had greater knowledge of reading strategies than their counterparts. No effects of vocabulary or decoding skills were found.

SES and ethnicity improved the model fit and were therefore retained in the model.

The classroom level predictors (the pretest, SES and ethnicity) did not have any effect, and no interactions were found involving the approach (DE or PI).

Further exploration of differential effects within the DE group and the PI group (Appendix Tables A1 and A2) revealed that in PI, only reading motivation ($b = 0.21, SE = 0.06, p < 0.01, f^2 = 0.05$) had a positive effect on knowledge of reading strategies. This indicates that in PI, more motivated students have better knowledge of reading strategies than less motivated students. In DE, only reading comprehension had a positive effect on knowledge of reading strategies ($b = 0.21, SE = 0.07, p < 0.01, f^2 = 0.06$). In DE, background variables with a significant effect were gender ($b = 0.28, SE = 0.11, p < 0.01, f^2 = 0.03$) and ethnicity ($b = 0.28, SE = 0.13, p < 0.05, f^2 = 0.00$). Furthermore, SES has a noticeable effect size on knowledge of reading strategies ($b = -0.30, SE = 0.16, p > 0.05, f^2 = 0.12$). All this means that in DE, students with good reading comprehension, girls and students with one or both non-native Dutch parents have greater knowledge of reading strategies than their counterparts.

The effects on reading motivation

The third research question could not be answered in the affirmative. As illustrated in Table 6, the educational approach did not affect reading motivation. Students in DE were not more motivated to read informative texts than students in PI.

Of the reading variables, reading comprehension ($b = 0.14, SE = 0.05, p < 0.01, f^2 = 0.02$) and knowledge of reading strategies ($b = 0.11, SE = 0.05, p < 0.01, f^2 = 0.02$) positively affected reading motivation.

Of the predictors, only decoding skills had a positive effect on reading motivation ($b = 0.17, SE = 0.04, p < 0.01, f^2 = 0.04$). Skilled readers were more motivated to read informative texts than less competent readers. Again, SES and ethnicity improved the fit of the model and were therefore retained. The classroom level predictors had no effect and did not improve the model fit.

Further exploration of both groups separately revealed some differential effects (Appendix Tables A1 and A2). In DE, reading comprehension had a positive effect on reading motivation ($b = 0.22, SE = 0.07, p < 0.01, f^2 = 0.04$). No significant effects were found for the other predictors. This means that in DE, students with good reading comprehension are more motivated to read than students with poor reading comprehension. In PI, knowledge of reading strategies ($b = 0.19, SE = 0.06, p < 0.01, f^2 = 0.04$), and decoding skills ($b = 0.22, SE = 0.06, p < 0.01, f^2 = 0.05$) positively affected reading motivation. This indicates that in PI, students with good decoding skills and good knowledge of reading strategies are more motivated than their counterparts. Furthermore, in PI, SES has a positive effect on reading motivation ($b = 0.33, SE = 0.15, p < 0.05, f^2 = 0.00$). This indicates that in PI, students with less educated parents are more motivated to read informative texts than students with highly educated parents.

Conclusion and discussion

The aim of the present study was to compare the effects of two approaches to reading comprehension instruction; i.e. the innovative meaning-oriented practice of DE and the more traditional practice
of PI. These approaches were compared in terms of reading comprehension, knowledge of reading strategies and reading motivation.

The results of the study revealed that there were no differences between students who were in DE and those who were in PI in reading comprehension, knowledge of reading strategies and reading motivation, after controlling for confounding variables. These findings indicate that both approaches had similar effects on each of the three outcome variables.

Nevertheless, some subtle distinctions were found when exploring the results for the two approaches separately. Concerning the reading variables (reading comprehension, knowledge of reading strategies and reading motivation), some mutual influences were found. In DE, more motivated readers and readers with more knowledge of reading strategies performed better in reading comprehension, which is in line with Guthrie's findings (Guthrie, Wigfield, and You 2012). Vice versa, readers with good comprehension had more knowledge of reading strategies and were more motivated for reading. In PI, more motivated readers had better knowledge of reading strategies, and vice versa, students with better knowledge of reading strategies were more motivated for reading informative texts. This is in line with earlier findings that good readers are more motivated and that motivated readers are better readers (Allington 2002).

A number of predictors influenced reading comprehension. In accordance with Babayiğit (2014), Snow (2002) and Van Elsäcker (2002), reading comprehension was predicted by vocabulary in both approaches. High non-verbal IQ was also an advantage for reading comprehension (Van Elsäcker 2002), although in a stronger way within the PI approach, than within the DE approach. Furthermore, in DE, students with better decoding skills had better reading comprehension. In PI, students with better decoding skills were more motivated to read informative texts, in line with Allington (2002). In agreement with the finding of Droop et al. (2012), in DE girls were found to have greater knowledge of reading strategies than boys.

The effects of the background variables SES and ethnicity were not clear. No significant effects on the outcome variables were found for either factor. Nevertheless, both variables improved the model fit in all analyses. In general, we may conclude that, for this sample, the educational approaches under investigation did not lead to significant differences in students’ performances in reading comprehension, knowledge of reading strategies or reading motivation.

Educational policy-makers may be reassured by our findings that the reading comprehension and knowledge of reading strategies that result from DE are as favourable as those that result from PI. However, as these findings contrast with the findings of Guthrie et al. (2004) and Guthrie and Klauda (2014) regarding CORI, they require further clarification. Whereas CORI was found to have a more positive effect on reading comprehension than traditional approaches, DE was not. The possible explanations for these findings are theoretical and practical in nature. A methodological explanation may be that we controlled for both a large number of confounding variables and contextual influences (i.e. multilevel classroom level influences). As such, false positive effects were maximally discarded as were the influences of the teachers and classroom populations. Teacher quality is one of the most critical variables in student achievement; 43% of the variance in student achievement can be attributed to teacher quality (Ferguson 1991 in: Snow 2002; 48). Not controlling for the teacher’s influence may distort the results of a study. Our teacher questionnaire was directed towards the implementation of the approaches and did not provide information about teacher quality or actual classroom practices. According to McKeown, Beck, and Blake (2009), it is possible that either some reading activities are more effective than others or that instructional activities matter less than simply the amount of time spent reading and the attention given to the text. Different aspects of DE and PI may contribute to the same outcomes. In PI, reading comprehension may be improved by a fixed amount of scheduled time and attention for reading tasks; whereas in DE, improvements in reading comprehension may be elicited by the focus on meaning. Following this line of reasoning, the focus on meaning applied in DE may lead, for example, to extensive classroom discussion and a reduction in the amount of time spent on the reading task itself (Allington 2002).
Strikingly, research on CORI has systematically reported positive effects on reading motivation compared to other approaches, whereas our investigation of DE did not find such effects. These findings are even more surprising given that CORI and DE share a number of basic tenets related to reading for meaning. A possible explanation may be related to the different contexts in which these studies were conducted. Whereas, DE is an existing practice, CORI is an intervention. Therefore, in contrast with DE, in the studies of CORI, the enhancement of motivation may have been novel to students and teachers and may have influenced the outcomes. In other intervention studies, motivational enhancement has also been found to affect reading motivation (Aarnoutse and Schellings 2003; Droop et al. 2012). Moreover, in this study, groups were studied in their natural context, in which many parameters cannot be controlled for as would be the case in an intervention.

Another explanation may be related to the construct of reading motivation itself, which requires further clarification (Schiefele et al. 2012), and the limited set of subscales used in this study. Only one out of three subscales of Guthrie’s MRQ (2010), curiosity, could be used because of low reliability, and four additional subscales tapping intrinsic motivation, self-efficacy, utility value and instrumental motivation. Moreover, in DE, reading is not the only activity that is used by students in the inquiry-oriented curriculum to answer questions about content. Children engage in classroom talks, listen to expert explanations, perform experiments or participate in field trips; these activities may motivate students to a greater extent than reading per se. In general, specific content may contribute to students’ motivation to read a specific text at a specific moment, but this does not necessarily lead to increased motivation for reading in general. Meaning is an important factor for reading, but reading is not the only way to construct meaning in DE.

A more practical explanation for the absence of differences in reading motivation may be our disregard for the fact that we do not know whether evidence from research in the US can be straightforwardly transferred to the Dutch educational system, which has its own specific characteristics. An indication of this non-transferability may be found in our questionnaire about the participating teachers’ reading practices. The responses revealed that the two approaches did not differ in terms of motivational aspects as much as we had expected they would. The teachers employing the PI approach did not differ significantly from those employing the DE approach in three of the six aspects of the questionnaire that involved motivation: interesting texts, self-efficacy and collaboration. It is possible that the two approaches did not differ from each other in these respects in practice as much as they did in theory. Modern reading programmes that are based on textbooks in the Netherlands have integrated these motivational aspects to a certain extent. However, this explanation is speculative because our data did not provide sufficient systematic information regarding the actual practices that teachers employed in the classrooms.

An extensive research base has provided evidence that knowledge of reading strategies is required to understand informative texts (NICHD 2000). The manner in which these strategies should be instructed is less evident. Our results do not support the prevailing model of PI as the best method for providing strategy instruction. Strategy instruction may be as effective if it is provided flexibly, i.e. if it is required for understanding as is the case in DE. In this respect, the present study contributes to the relatively small number of studies of content-based approaches (McKeown, Beck, and Blake 2009).

Several limitations of this research should be noted. First, the design consisted of a limited number of schools (12 schools in each group), which may have resulted in limited power and prevented the observation of the intended effect, which may exist in reality (i.e. within the total populations of schools and students). Second, the participating schools were not selected randomly. The number of schools that have adopted and implemented the DE concept in the upper grades is limited, and because we wanted the sample populations to vary in terms of ethnicity and SES, only a small number of DE schools met the criteria for selection. The small number of students from low SES families and ethnic minorities may be the reason why the influence of parental education could not be established or further explained.

Finally, it is a limitation of the present research that the examinations of the teachers’ instructional practices within both approaches were based on self-reports. DE schools may be heterogeneous as they
may implement different procedures for teaching reading for meaning. Systematic classroom observations would have been more reliable; however, such observations would have entailed insurmountable difficulties in terms of time and logistics. Another possibility would have been to ask students for their perception of the classroom practices as was done by Guthrie and Klauda (2014).

Despite these limitations, it may be cautiously concluded that the innovative DE approach produces results that are as good as those produced by the reading approach of PI. However, we were not able to confirm the assumption that the reading approach of DE adds value; i.e. we were unable to determine whether DE with its focus on reading for meaning, motivated students to read informative texts. Further research is needed to gain more insight into the added value of engagement in reading in relation to the other, further-reaching, goals of DE of identity development, agency and citizenship.

Note

1. In SPSS, two estimation methods are available: restricted maximum likelihood (RML) and ML. The latter estimation method is preferable when comparing the fit indices (−2 log likelihood) of two models that differ only in the numbers of fixed-effect parameters as was the case in the present study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on Contributors

Yvonne van Rijk is a teacher trainer and researcher at the Faculty of Education at Amsterdam University of Applied Sciences. Her interests focus on language acquisition and education from a sociocultural perspective. She is writing a dissertation on the teaching of meaningful reading in primary education.

Langha de Mey is an assistant professor in Methodology at the Department of Research and Theory of Education at VU University Amsterdam. His research focus is on quantitative methodology and data-analysis.

Dorian de Haan is a professor in Developmental Education at the Inholland University of Applied Sciences and an assistant professor at the Department of Developmental Psychology at Utrecht University. Her research focus is on language acquisition and education in multi-ethnic institutions of childcare and primary school.

Bert van Oers is a professor in Cultural-Historical Theory of Education at VU University Amsterdam (Faculty Psychology and Education). Since the 1980s, he is involved in the elaboration, implementation and evaluation of the Developmental Education Concept in the Netherlands. His main topics and publications are related to this project.

Monique Volman is Full Professor of Education and Director of the 'Educational Sciences’ programme at the Research Institute of Child Development and Education of the University of Amsterdam. Main areas in her research are learning environments for meaningful learning, diversity and the use of ICT in education.

ORCID

Bert van Oers <http://orcid.org/0000-0001-9371-9126>

References


Babayiğit, S. 2014. “The Role of Oral Language Skills in Reading and Listening Comprehension of Text: A Comparison of Monolingual (L1) and Bilingual (L2) Speakers of English Language.” Journal of Research in Reading 37 (S1), S22–S47.


## Appendix

Table A1. Final multilevel models of reading comprehension, reading motivation and knowledge of reading strategies of the DE approach.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Comprehension</th>
<th>Strategies</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$</td>
<td>SE</td>
<td>$\hat{f}^2$</td>
</tr>
<tr>
<td>Intercept</td>
<td>$-0.00$</td>
<td>(0.08)</td>
<td></td>
</tr>
<tr>
<td><strong>Level 1 fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>0.15**</td>
<td>(0.06)</td>
<td>0.04</td>
</tr>
<tr>
<td>Comprehension</td>
<td>$X$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivation</td>
<td>0.10*</td>
<td>(0.05)</td>
<td>0.01</td>
</tr>
<tr>
<td>Strategies</td>
<td>0.14**</td>
<td>(0.05)</td>
<td>0.03</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.21**</td>
<td>(0.06)</td>
<td>0.08</td>
</tr>
<tr>
<td>Decoding skills</td>
<td>0.13*</td>
<td>(0.05)</td>
<td>0.05</td>
</tr>
<tr>
<td>Non-verbal IQ</td>
<td>0.12*</td>
<td>(0.05)</td>
<td>0.03</td>
</tr>
<tr>
<td>SES ($0 = \text{low}$)</td>
<td>$-0.13$</td>
<td>(0.14)</td>
<td>0.28</td>
</tr>
<tr>
<td>Ethnicity ($0 = \text{not Dutch}$)</td>
<td>$-0.18$</td>
<td>(0.13)</td>
<td>0.03</td>
</tr>
<tr>
<td>Gender ($0 = \text{girl}$)</td>
<td>$0.28**$</td>
<td>(0.11)</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>Level 2 fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest classroom</td>
<td>0.17**</td>
<td>(0.06)</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Random effects**

<table>
<thead>
<tr>
<th></th>
<th>$\hat{f}^2$</th>
<th></th>
<th></th>
<th>$\hat{f}^2$</th>
<th></th>
<th></th>
<th>$\hat{f}^2$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual</td>
<td>0.36**</td>
<td>(0.04)</td>
<td></td>
<td>0.52**</td>
<td>(0.05)</td>
<td></td>
<td>0.63**</td>
<td>(0.07)</td>
<td></td>
</tr>
<tr>
<td>Class/school</td>
<td>0.03</td>
<td>(0.02)</td>
<td></td>
<td>0.01</td>
<td>(0.01)</td>
<td></td>
<td>0.03</td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>ICC</td>
<td>0.08</td>
<td></td>
<td>0.02</td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-2\log$ likelihood</td>
<td>350</td>
<td>421</td>
<td>466</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$ difference test</td>
<td>$\chi^2 (9) = 289**$</td>
<td></td>
<td></td>
<td>$\chi^2 (5) = 223**$</td>
<td></td>
<td></td>
<td>$\chi^2 (3) = 205**$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Blanks: no improvement of the model fit, and no significant effect, the variable was removed from the model. ** ($p < 0.01$); * ($p < 0.05$).

SES represents a low level of parental education.

Ethnicity represents a home language other than Dutch, or a mix of Dutch and another language.

Level 1 is the student level, and level 2 is the classroom level.

Pretest classroom was the first measurement occasion of the outcome variables at the classroom level.
Table A2. Final multilevel models of reading comprehension, reading motivation, and knowledge of reading strategies of the PI approach

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Comprehension</th>
<th>Strategies</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(b)</td>
<td>SE</td>
<td>(f^2)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.02</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>Level 1 Fixed effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>0.23**</td>
<td>(0.05)</td>
<td>0.08</td>
</tr>
<tr>
<td>Comprehension</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivation</td>
<td>0.05</td>
<td>(0.05)</td>
<td>0.00</td>
</tr>
<tr>
<td>Strategies</td>
<td>0.02</td>
<td>(0.05)</td>
<td>0.00</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.31**</td>
<td>(0.06)</td>
<td>0.11</td>
</tr>
<tr>
<td>Decoding skills</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-verbal IQ</td>
<td>0.35**</td>
<td>(0.06)</td>
<td>0.11</td>
</tr>
<tr>
<td>SES (0 = low)</td>
<td>-0.19</td>
<td>(0.13)</td>
<td>0.00</td>
</tr>
<tr>
<td>Ethnicity (0 = not Dutch)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (0 = girl)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 2 Fixed effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest classroom</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>0.46**</td>
<td>(0.04)</td>
<td>0.75**</td>
</tr>
<tr>
<td>Class / school</td>
<td>0.02</td>
<td>(0.02)</td>
<td>0.10</td>
</tr>
<tr>
<td>ICC</td>
<td>0.04</td>
<td></td>
<td>0.12</td>
</tr>
<tr>
<td>(-2)log likelihood</td>
<td>514</td>
<td></td>
<td>651</td>
</tr>
<tr>
<td>(\chi^2) difference test</td>
<td>(\chi^2 (6) = 220^{**})</td>
<td></td>
<td>(\chi^2 (5) = 123^{**})</td>
</tr>
</tbody>
</table>

Notes: Blanks: no improvement of the model fit, and no significant effect, the variable was removed from the model.  
** \((p < 0.01)\); * \((p < 0.05)\).  
SES represents a low level of parental education.  
Ethnicity represents a home language other than Dutch, or a mix of Dutch and another language.  
Level 1 is the student level, and level 2 is the classroom level.  
Pretest classroom was the first measurement occasion of the outcome variables at the classroom level.