Reading and Writing (2023) 36:2549–2575
https://doi.org/10.1007/s11145-022-10359-2

Does vocabulary knowledge matter in the effectiveness of instructing reading strategies? Differential responses from adolescents with low academic achievement on growth in reading comprehension

M. Okkinga[1] · A. J. S. van Gelderen[1,2] · E. van Schooten[1,3] · R. van Steensel[3,4] · P. J. C. Sleegers[5]

Accepted: 9 September 2022 / Published online: 30 December 2022

Abstract

Prior studies suggest that teaching reading strategies promotes reading comprehension in adolescents who have difficulties with reading comprehension, yet the results of those studies are mixed. Individual differences in students' vocabulary knowledge may explain these mixed results. This article examines to what extent vocabulary knowledge influences the effect of a two-year intervention program focused on teaching reading strategies to adolescents with low academic achievement in the Netherlands. We hypothesized that students (N = 310) with different levels of vocabulary knowledge would respond differently to the treatment, given that vocabulary knowledge is an important factor in reading comprehension. Results showed that vocabulary knowledge moderated the effect of the treatment, suggesting that low vocabulary knowledge negatively affected the impact of an intervention focused on reading strategies. Vocabulary knowledge, thus, emerges as a prerequisite for the successful leveraging of a reading strategy intervention. Students with low vocabulary knowledge may experience cognitive overload when attempting to apply newly learned reading strategies while simultaneously trying to find out the meaning of multiple unfamiliar words needed for successful application of reading strategies.

Keywords Adolescents with low academic achievement · Reading comprehension · Reading strategies · Vocabulary knowledge · Moderating effect

Many adolescents struggle with reading comprehension (e.g., Organisation for Economic Co-operation and Development [OECD], 2003; OECD, 2014). Since it is a fundamental skill in many school subjects, poor reading comprehension has serious

Extended author information available on the last page of the article

implications for students' educational success and, consequently, for their later societal careers. Since the 1980's, reading programs have focused on teaching reading strategies in order to foster reading comprehension (Pressley & Afflerbach, 1995; Raphael et al., 2009). However, results of research in teaching reading strategies to adolescents with low academic achievement are mixed (Edmonds et al., 2009; Fogarty et al., 2014; Okkinga et al., 2021; Simmons et al., 2014; Vaughn, 2013). A possible explanation for the different results is that these studies do not take individual differences in students' skills into account. In the present study we analyzed whether vocabulary knowledge moderated the effect of instruction in reading strategies on reading comprehension in adolescents with low academic achievement.

Reading comprehension and vocabulary knowledge

Reading comprehension is a complex process involving several sets of knowledge and skills. Most models of reading comprehension distinguish between lower and higher order skills that interact in the process of creating a mental representation of a text, or a situation model (Kintsch, 1998). To construct a coherent mental representation, the information in the text is integrated with the readers' background knowledge. Lower order skills include letter and word recognition (Perfetti et al., 2005), while higher order skills refer to the ability to give meaning to words and sentences, make inferences and representations of paragraphs or a text as a whole (Aarnoutse & Van Leeuwe, 1988). Vocabulary knowledge is a basic component of these higher order skills of reading comprehension.

It is safe to assume that if many words of a text are not properly understood, it becomes difficult to comprehend the text (Torgesen, 2000). This assumption is supported by ample empirical evidence in which a strong relationship between vocabulary knowledge and reading comprehension is found (Ahmed et al., 2016; Ouellette & Beers, 2010; Tilstra et al., 2009; Trapman et al., 2014; Van Gelderen et al., 2004; Gelderen et al., 2007; Van Steensel et al., 2016; Verhoeven & Van Leeuwe, 2008).

Instructing reading strategies to foster reading comprehension

A reading strategy is a mental tool a reader uses purposefully to monitor, repair, or bolster comprehension (Afflerbach & Cho, 2009). The use of reading strategies is a deliberate and goal-directed attempt to construct meaning from text (Afflerbach et al., 2008). In the literature, metacognitive and cognitive strategies that aid the process of reading are distinguished (Dole et al., 2009). Researchers have suggested many different strategies (Pressley & Afflerbach, 1995). Important strategies involve setting explicit reading goals, activating relevant background knowledge, allocating attention to major content while ignoring irrelevant details, evaluating the validity of text content, comprehension monitoring, and making and testing interpretations, predictions, and drawing conclusions (Palincsar & Brown, 1984).

A widely used method of instructing, teaching, and guiding adolescents with poor reading skills in the use of reading strategies to foster reading comprehension

is reciprocal teaching (Palincsar & Brown, 1984; Palincsar et al., 1987). Reciprocal teaching consists of a set of three principles: (a) teaching comprehension-fostering reading strategies (b) expert modeling, scaffolding and fading; and (c) students taking turns in practicing reading strategies and discussing with other students. The method includes the teacher explicitly modeling the use of reading strategies during the start of reciprocal teaching (Rosenshine & Meister, 1994) as well as scaffolding the application of reading strategies within the groups of students working together. During this process, students become increasingly more capable of regulating their own reading process and the role of the teacher gradually fades. Many studies have demonstrated positive effects of this approach (Kelly, Moore, & Tuck, 2011; Palincsar & Brown 1984; Rosenshine & Meister, 1994; Spörer et al., 2009; Webb et al., 2019). However, there are also indications that the approach of teaching reading strategies based on reciprocal teaching is not always successful in improving adolescents' reading comprehension (Edmonds et al., 2009; Fogarty et al., 2014; Muijselaar et al., 2018; Okkinga et al., 2021; Simmons et al., 2014; Vaughn, 2013). An important factor that might explain these mixed results is the role of vocabulary knowledge in reading comprehension and strategy use.

Most interventions aimed at instructing reading strategies to foster reading comprehension in adolescents with low academic achievement are based on research that is focused on characteristics of the reading process of good readers (Pressley & Afflerbach, 1995; Raphael et al., 2009). The idea is that if students who struggle with reading comprehension are taught the reading strategies used by skilled readers, their reading behavior will change and their comprehension of texts will improve as they learn to implement these strategies. However, this line of reasoning overlooks other important skills that support reading comprehension. In particular, readers who struggle with reading comprehension differ from skilled readers not only in their use of reading strategies, but also in their vocabulary knowledge. As has been widely documented, vocabulary knowledge explains adolescents' reading comprehension to a considerable degree, even within the group of adolescents with low academic achievement (Trapman et al., 2014; Trapman, Van Gelderen, Van Schooten, & Hulstijn, 2017; Van Gelderen et al., 2004, 2007; Van Steensel et al., 2016). Adolescents with little vocabulary knowledge may be less able to profit from strategies derived from the good readers' reading processes. These students might be also less likely to profit from instruction in reading strategies than students who struggle with reading comprehension but with a larger vocabulary.

For example, a strategy such as monitoring comprehension of sentences depends for a large part on the degree of accuracy of knowledge of the words used in these sentences. From the literature about reading comprehension, it is known that the reading process including the conscious use of strategies for comprehension is cognitively highly demanding (Kendeou et al., 2014). Perfetti and Hart (2002) point to the role of lexical quality in the reading process and to the consequences for readers suffering from low lexical quality. Lexical quality pertains to the connections made between the orthographic, phonological and semantic properties of words. For readers who have a low lexical quality, which may pertain to each of these properties of vocabulary knowledge, overcoming word-level comprehension problems can be cognitively demanding. Therefore, it is plausible that poor readers have to use their working

memory primarily for word-level issues and therefore lack working memory space for the application of more general strategies for comprehension. Adolescents suffering from insufficient vocabulary knowledge will therefore be at a disadvantage in the use of strategies for monitoring comprehension, one of the most important activities in the approach of reciprocal teaching practiced by Palincsar & Brown (1984).

Following this line of reasoning, we hypothesize that adolescents with low academic achievement differing in their level of vocabulary knowledge will not benefit similarly from an intervention aimed at instructing reading strategies to foster reading comprehension. In other words, these adolescents may differ in vocabulary knowledge and this may influence their response to a reciprocal teaching intervention directed at the improvement of reading strategies. Specifically, we hypothesize that students with a smaller vocabulary will benefit less from the intervention than students with a larger vocabulary.

The present study

In the present study, reciprocal teaching was used in a two-year intervention to improve reading comprehension of adolescents with low academic achievement, defined as students placed in the two lowest tracks of the Dutch secondary education system on the basis of a general academic attainment test (see Sample). Previously, we found small effects of implementation quality of the intervention on students' growth in reading comprehension between students in the experimental and control condition, but no overall main effect (Okkinga et al., 2018, 2021). It is possible that students' vocabulary knowledge moderated the effect of the intervention. As discussed above, differences in vocabulary knowledge can be decisive for reading comprehension and for the application of reading strategies and therefore may moderate the degree of growth in reading comprehension.

For that reason, we examine the moderating role of vocabulary knowledge in the effects of the intervention on growth in reading comprehension.

We will answer the following research question:

Is the effect of a reading strategy intervention to improve reading comprehension in adolescents with low academic achievement moderated by vocabulary knowledge?

Method

Design

A two-year longitudinal design with a cluster randomized controlled trial was used in this study (Shadish et al., 2002). Randomization took place at the class level. At every participating school two classes, each with their own Dutch language teacher, took part in the study. The dependent variable, reading comprehension, was mea-

sured at four time points. We included control variables on the student level (gender, IQ, language background, metacognitive knowledge and age). Finally, vocabulary knowledge was included as a moderator variable.

Gender was included as a control variable because girls generally show better reading skills than boys (Logan & Johnston, 2009; Schaffner et al., 2016). Higher scores on IQ and metacognitive knowledge (defined as knowledge of text characteristics and reading strategies) correlate with higher scores on reading comprehension and, therefore, also IQ and metacognitive knowledge were included as control variables (Just & Carpenter, 1976, 2004; LaBerge & Samuels, 1974; Rumelhart, 2004; Samuels, 2004; Ouellette & Beers, 2010; Van Gelderen et al., 2004; Gelderen et al., 2007; Verhoeven & Van Leeuwe, 2008; Trapman et al., 2014). Since empirical evidence suggests that factors contributing to reading comprehension are different for adolescents with low academic achievement who mainly speak another language than Dutch at home compared to students who speak mainly Dutch at home (Trapman, 2015), language background was also included as a control variable. In addition, age was included as a control variable.

Sample

Our study focused on adolescents with low academic achievement. Our operationalization of low achievement was based on educational track. The Netherlands has a tracked system of secondary education. After primary school, students are placed in one of three tracks—prevocational secondary education, senior general secondary education, or pre-university education— on the basis of their scores on a general academic attainment test and their educational performance as assessed by their primary school teachers (Ministry of Education, Culture, & Science, 2006)[1]. Since students in prevocational education tend to have poor reading skills (Dutch Education Inspectorate, 2008; Gille et al., 2010), we selected our sample from schools offering this type of education.

We recruited schools in two ways. First, we contacted schools that had participated in a previous study. Second, we contacted schools via a digital community of Dutch language teachers. Schools had to meet the following five criteria:

- Willingness to participate in a treatment study.
- They had (at least) two seventh grade classes.
- Each class had its own Dutch language teacher.
- The teachers were prepared to take part in the randomization procedure, implying that (a) if their class was assigned to the treatment condition, they were prepared to take part in our training and coaching program and to weekly give the experimental lessons; and (b) if their class was assigned to the control condition, they were prepared to not use our program nor discuss its contents with the colleague in the treatment condition.

[1] The prevocational track is subdivided in three types. We selected our sample from the two lowest of these, representing about the 30% lowest scoring on the general attainment test.

- Control teachers were requested to use their regular reading program during the language classes.

Ten different schools in different parts of the Netherlands were willing to participate. Within each school, two teachers volunteered. Randomization was done at the class level within each school, resulting in a total of ten experimental and ten control classes. At the start of the study, these classes comprised 369 students, of which 189 were in the treatment condition (51%) and 180 in the control condition (49%). The students' mean age was 13.01 years ($SD = 0.52$) at the start of the project. The mean ages of the students in the two conditions did not differ significantly ($t(366) = -1.27$, $p = .20$). There were relatively more girls in the sample ($n = 200$; 54%) than boys ($n = 169$; 46%), with relatively more girls than boys (59 vs. 41%) in the treatment condition. The distribution in the control condition, however, was more balanced (49 vs. 51%). The difference in distribution between the two conditions was statistically significant (χ^2 (1) = 3.99, $p = .046$). Most of the students had a Dutch language background (84.2%), while a minority more often spoke another language than Dutch at home (15.8%). The distribution of language background across conditions did not differ statistically (χ^2 (2) = 0.024, $p = .988$).

More female than male teachers participated in the study ($n = 15$ vs. $n = 5$). There were two male teachers in the treatment group and three in the control group. The mean age of the teachers was 46.40 years ($SD = 11.12$). On average they had 13.50 ($SD = 13.73$, $min = 1$, $max = 38$) years of teaching experience in secondary education. No significant differences were found between the conditions on age (t (14) = −0.45, $p = .66$) and years of experience (t (14) = 0.053, $p = .96$).

Attrition and teacher replacements

One experimental class dropped out of the study after one year, because the teacher was replaced by another teacher who was not willing to continue the cooperation. The data of this class (24 students) were therefore not included in this analysis. In two other experimental classes replacement of teachers took place. Both teachers got seriously ill, one after half a year, the other after the first year. They were replaced by other teachers who continued the cooperation, but it took the schools several weeks to find the replacement. In the control condition, three teachers were immediately replaced, due to illness, a new job or because the teacher had to teach other classes. One of these replacements occurred at the end of the second year, one after one year, the third after half a year. In order to account for these replacements and missed classes, two control variables were added to our analysis: teacher replacement (0, 1) and cancelled classes (0, 1).

There was considerable attrition among the students. From a total of 369 students at the start of the project, 44 students changed schools, of which 19 students were in the treatment condition and 25 in the control condition. Seven students ($n_{\text{treatment}} = 4$; $n_{\text{control}} = 3$) switched classes within their school. Three students were ill for a longer period of time during the study, of which two were in the treatment condition. The frequency distribution of these categories (students staying, changing schools,

switching classes, and illness) across the treatment and control condition did not differ significantly, $\chi^2(3)=4.78$, $p=.19$. In addition, we excluded students ($n=5$) who were not at least present during the first two test sessions (i.e. the first year of the intervention). A total of 310 students were included in the analyses, of which 165 students in the treatment condition and 145 students in the control condition.

Treatment

Our intervention consisted of the training of five strategies that were shown to be related to reading comprehension in previous research (Dole et al., 1991; Palincsar & Brown, 1984; Pressley & Afflerbach, 1995; Van Silfhout et al., 2014):

1. *Predicting.* On the basis of text features such as title, subheadings, and pictures, students are instructed to make predictions about text content before reading, and to check their predictions while reading.
2. *Summarizing.* Students are instructed to summarize sections of text, encouraging them to focus on main ideas and ignore irrelevant details as well as to check their understanding of the text so far.
3. *Self-questioning.* Students are instructed to generate questions about the text being read, helping them to focus on main ideas as well as to monitor understanding.
4. *Clarifying.* When confronted with a word or passage they do not understand, students are instructed to reread, read ahead, or, in the case of an unknown word, analyze it, and see whether its meaning can be inferred by looking at parts of the word.
5. *Interpreting cohesive ties*. Students are instructed to look for relationships between sentences or paragraphs that are connected, for instance, by using 'signal words' (different types of connectives).

These reading strategies were taught in the context of an existing program called "Nieuwsbegrip"®, developed by the CED Group in Rotterdam ("Newswise"). Lessons were developed weekly by a team of developers at the CED Group and were based on recent news articles (i.e., texts that had been issued the week before) about subjects related to students' everyday life (e.g., sugar in energy drinks, abdication of the Dutch queen, or 20 years of text messaging). The use of topical, interesting texts was assumed to increase students' task motivation (Guthrie & Wigfield, 2000; Schiefele, 1999). The teachers could download the lessons from the program website (www.nieuwsbegrip.nl) and were made available every Monday evening of the week.

Lessons were provided in sequences, each consisting of six weekly lessons (approximately 45 min per lesson). In each of the first five lessons, the focus was on one reading strategy that was practiced in an assignment provided on a worksheet. In addition, students could work on other assignments (i.e., answering questions about the text) on the worksheet. The total duration of the intervention was 70 lessons divided over two school years. The lessons started after the autumn holiday, and lasted until the summer holidays of the second intervention year. The actual number of lessons provided varied, since teachers were occasionally ill or schoolwide events

Table 1 Examples of assignments for each reading strategy

Strategy	Example
Predicting	This text has five subheadings. Write down for each subheading (a) which thoughts it evokes and (b) what you already know about the subject addressed in the subheading.
Summarizing	Read the text. Read paragraph by paragraph and underline in each paragraph the most important information. For each paragraph, write one or two sentences summarizing it. Use the words you underlined.
Self-questioning	Read the text. Note at least five questions that spring to mind while reading.
Clarifying	Search the text for difficult words. Try to uncover their meaning using these hints: (a) reread the previous piece of text or read on, (b) look at the illustrations in the text, (c) look at the word: you might know part of the word, (d) sometimes you have to use your own knowledge to figure out word meanings, or (e) use a dictionary.
Interpreting cohesive ties	Read the text. Underline the signal words. Answer the questions, while noting the signal words: • Which contrast is explained in lines 16–17? [signal word=however] • Why are energy boosters unfit as sports drinks? [signal word=hence]

took place. In classes where more than 6 intervention lessons were cancelled, this was taken into account as a control variable (see supplementary material Appendix B).

Students were trained in each of the five strategies several times during the year. This cyclical approach was assumed to result in the consolidation of strategy use. In the final lesson of each sequence all strategies were practiced simultaneously. The idea behind this was that students have to be able to apply all strategies together during the reading process, selecting an appropriate strategy at the right moment. Table 1 provides an example of an assignment for each reading strategy.

The didactic approach underlying the program "Newswise" is reciprocal teaching (Palincsar & Brown, 1984). As mentioned earlier, reciprocal teaching is a widely used method of instructing and guiding learners in reading comprehension. It consists of a set of three related instructional principles: (a) direct teaching of comprehension-fostering reading strategies, including predicting, question-generating, summarizing, and clarifying; (b) expert modeling, scaffolding and fading; and (c) students practicing and discussing reading strategies with other students, guided and coached by the teacher. Reciprocal teaching assumes a gradual shift of responsibility for the learning process from teacher to student, which includes the teacher explicitly modeling the use of reading strategies (Rosenhine & Meister, 1994) as well as scaffolding the application of reading strategies within the groups of students working together. It is assumed that by gradually fading teacher support, students become increasingly more capable of regulating their own reading process.

Treatment teachers took part in a training and coaching program that was conducted by teacher trainers from the Rotterdam University of Applied Sciences, who had, in turn, been trained by Authors 1, 2, and 4. In the first phase (October 2011-January 2012), teachers participated in three one-hour training sessions, covering the didactic principles of the programme. In the second phase (February 2012-May 2013), teachers participated in six coaching sessions; three coaching sessions during February-June 2012 and three coaching sessions during September 2012 – May 2013. A coaching session involved a classroom observation conducted by the trainer during an intervention lesson, followed by a feedback meeting of approximately twenty minutes on the same day (See for detailed information about treatment implementation and fidelity: Okkinga et al., 2021).

Control classes

Control classes were "business as usual". Teachers in the control classes used the regular textbook for Dutch language teaching that was used in their school. Among the schools in the study, three different language textbooks were used. The textbooks and their teacher manuals were analyzed according to the three principles of instructional strategies in the treatment condition: instruction of reading strategies, modeling, and group work. Attention was given to reading strategies in all three textbooks. However, not all strategies that were covered in the treatment condition were also covered in the control textbooks. Reading strategies that were referred to were: predicting, clarifying, and attention to cohesive ties. Self-questioning did not occur and little attention was given to summarizing.

No attention was given to modeling by teachers or students in the teacher manuals of the control classes. Almost all of the assignments were individual and there were only a few instances where students were instructed to work together on an assignment.

Treatment fidelity

We conducted classroom observations in both the experimental and control conditions twice each year, resulting in a total of four observations for each class. Our aim was to examine (a) whether the treatment teachers provided the lessons in the way we instructed during the training and coaching program and (b) whether the control teachers applied the three treatment principles, even though they were not trained to do so. We used four-point scales (0–3) to assess the treatment fidelity of the main three elements of the intervention: Whole-class teaching of reading strategies, teacher and student modeling, and group work (see for more information Okkinga et al., 2018). The four observations per class were averaged to create one score for each element of treatment fidelity (Table 2). As can be seen in Table 2, the treatment classes scored higher on all three elements.

Inter-rater reliability was calculated by means of observed agreement between two observers. In total, 30 from a total of 76 classroom observations were performed

Table 2 Descriptives Treatment fidelity

Variable	Treatment ($n=8$)	Control ($n=10$)		
	Mean (SD)	Mean (SD)	t-value	p-value
Strategy-instruction	1.81 (0.80)	1.13 (0.65)	2.02	0.061
Modeling	0.84 (0.65)	0.33 (0.35)	2.15	0.047
Group work	2.00 (1.14)	0.30 (0.33)	4.51	<0.001

Note. Scoring between the three variables cannot be compared one-on-one. The meaning of the scoring (0–3) is different for each variable.

independently by two coders. Across these 30 observations, 93.89% observed agreement was obtained.

Measures

Reading comprehension

Reading comprehension was measured by means of the SALT-reading, a test that was validated for use among adolescents with low academic achievement (Van Steensel et al., 2013). The SALT-reading comprises eight tasks, each consisting of one or two texts and comprehension questions about those texts. The texts cover different genres (narrative, expository, argumentative, and instructive) and were selected from media students come across regularly in their daily lives: (school) books, newspapers, magazines, and official documents (such as regulations in a youth hostel). The eight tasks comprised a total of 59 test items, that were divided into three categories: items requiring students to retrieve relevant details from the text, items requiring students to make inferences on a local level (e.g. cause-effect relationships between sentences), and items requiring students to show their understanding of the macro structure of the text (e.g., by inferring the main idea of the text or the intention of the author). The test consisted mainly of multiple choice questions but contained also five open-ended questions. The SALT-reading was administered at four time points (See Design). The Cronbach's alpha coefficients were 0.82, 0.83, 0.82, and 0.85 respectively.

Vocabulary knowledge

Vocabulary knowledge was assessed with a 73-item multiple-choice test, measuring the knowledge of nouns, verbs, adjectives, and adverbs belonging to the 23,000 words in a dictionary for junior high school students (see Hazenberg & Hulstijn 1996, for details). Each item consists of a neutral carrier sentence with a bold-faced target word and four answer options, one of which represents a correct synonym. The vocabulary knowledge test was administered two times (at pretest and at the end of the first school year). The Cronbach's alpha coefficients were 0.86 and 0.85, respectively. As there was a substantial correlation between the repeated measure for vocabulary ($r=.72$) and there was no relation between the test items and the contents of the lessons, we decided to use both sources of information in our measure for

vocabulary knowledge to increase the reliability of the measurement. Therefore, the average of the repeated measure was used as a measure for vocabulary knowledge.

Metacognitive knowledge

Metacognitive knowledge was assessed with Trapman's and colleagues' (2014) adapted version of the metacognitive knowledge test directed at adolescents (grades 8 to 10) constructed by Van Gelderen et al. (2003) and Van Gelderen et al. (2007).

Items consisted of correct or incorrect statements about text characteristics and reading and writing strategies. Students had to indicate whether they agree or disagree with each statement. An example of an incorrect statement about text characteristics is 'The order in which you present the information in your text is usually not relevant'. An example of a correct statement about text characteristics is 'Sometimes you need to know more than what is in the text to understand the text properly'. The test consisted of 45 items and was administered at the end of the first school year. The Cronbach's alpha coefficient was 0.51. Although this indicates a rather low level of reliability (Field, 2009), the measure predicted significant variance in reading comprehension in previous research (Trapman et al., 2017). Therefore, we maintained the measure in the analyses.

Nonverbal IQ

Intellectual ability was measured by administering the Raven Progressive Matrices at pretest. The test consists of 60 items, divided into 5 sets of 12 items. Each item represents a logical reasoning puzzle. The items become more difficult within a set and the sets become increasingly difficult as well (Raven et al., 1998). A pilot of the test for our target group showed that the last set was too difficult for these students. Therefore, we administered only four sets. The Cronbach's alpha coefficient was 0.82.

Language background

Language background was assessed with a questionnaire asking students which language they predominantly speak with their primary caregiver. There were three options: (1) the student speaks predominantly Dutch with his or her primary caregiver, (2) the student speaks predominantly a different language than Dutch with his or her primary caregiver, and (3) the student speaks as much Dutch as another language with his or her primary caregiver. Language background was operationalized as an ordinal variable with students who mainly spoke Dutch with their primary caregiver coded as 0 and students who spoke a different language than Dutch with their primary caregiver coded as 2, while students who spoke as much Dutch as a different language were coded as 1.

Analyses

Missing data treatment

After the collection of all data, the dataset contained data of 310 students. To prevent loss of students in the multilevel analyses, occasional missing items were coded as incorrect and for other instances (students missing a whole page) missing data were imputed for independent variables (never exceeding 7% of the cases per item), using the EM procedure from SPSS missing value analysis.

Procedure of multilevel analyses

Repeated measures multilevel regression analyses were performed to account for the hierarchical structure of the data (using MLwiN 2.16; Rasbash et al., 2009). The time variable 'Occasion' (variance within students across times of measurement) was defined in months; with the first measurement of reading comprehension at month zero, and subsequent measurements at months 9, 12, and 22, respectively. These months correspond to the following time points: September 2011, June 2012, September 2012, and June 2013. Dichotomous independent class and student variables were always scored 0 and 1, all continuous independent variables (IQ, age, vocabulary and metacognitive knowledge) were centered around the grand mean before adding them to the model (Hox, 2010).

First, we tested whether adding a class or school variance level to the model significantly improved model fit. Levels significantly improving model fit were added to the model. Second, we tested whether a model with random slopes both at the student or class level improved model fit, as is recommended by Hox (2010). Results for these two steps are presented in Appendix A. Third, we added the class level variables 'teacher replacement' and 'cancelled classes' to check whether we should correct for these variables (see supplementary material Appendix B). Fourth, we tested whether the student-level control variables gender, IQ, age, language spoken at home with primary caregiver, and metacognitive knowledge significantly improved model fit. In addition, vocabulary knowledge was included as a predictor, because this was necessary for the later testing of its moderating effect (see Table 5). Fifth, we tested whether the treatment had a significant effect on growth in reading comprehension by testing the interaction between treatment and occasion. Finally, we checked whether vocabulary moderated the treatment effect (see Table 6) by estimating the three-way interaction between occasion, treatment and vocabulary knowledge.

Testing of significance

The number of levels needed in the analyses was tested by comparing nested models with Chi-square significance tests (Hox, 2010). Significance of control variables and predictors were tested both with Wald-tests (coefficient divided by the standard error) and by means of comparing nested models (with and without the control variables or

predictors) with a Chi-square test[2]. Regression coefficients for class level variables were tested with number of classes as sample size (df = number of classes – number of predictors – 1) (Hox, 2010). Effect sizes were estimated by comparing the variances at different levels as well as the total variances of nested models.

Results

Descriptive statistics

Table 3 shows the mean student scores for all student-level variables (the four repeated measurements of reading comprehension and test scores for IQ, vocabulary knowledge, and metacognitive knowledge). The only student level variable showing a significant difference between treatment and control group (t-tests) is metacognitive knowledge. In Table 4 Pearson correlations are presented between all student-level variables. All correlations are significant at $p < .01$. The correlations do not indicate high multicollinearity, as all correlations are below 0.8 (Field, 2009).

Table 3 Descriptives student-level variables

Variable	Treatment ($n = 165$)	Control ($n = 145$)	
	Mean (*SD*)	Mean (*SD*)	t-value
1. Reading comprehension (time 1)	34.96 (6.98)	34.59 (8.50)	0.41
2. Reading comprehension (time 2)	37.49 (6.92)	36.47 (8.81)	1.12
3. Reading comprehension (time 3)	36.45 (7.31)	36.94 (8.33)	0.51
4. Reading comprehension (time 4)	37.35 (8.51)	39.24 (8.60)	1.75
5. Vocabulary knowledge	49.30 (7.42)	49.37 (7.69)	0.08
6. IQ	35.96 (5.24)	35.09 (5.27)	1.45
7. Metacognitive knowledge	26.92 (4.36)	25.59 (4.46)	2.66*

*$p = .008$

Table 4 Correlation matrix student-level variables

Variable *N*	1. RC1 310	2. RC2 310	3. RC3 259	4. RC4 250	5. VK 310	6. IQ 310	7. MK 310
1. Reading comprehension time 1 (RC1)	-	0.68	0.62	0.62	0.59	0.29	0.35
2. Reading comprehension time 2 (RC2)		-	0.75	0.70	0.60	0.28	0.42
3. Reading comprehension time 3 (RC3)			-	0.73	0.56	0.32	0.35
4. Reading comprehension time 4 (RC4)				-	0.58	0.23	0.32
5. Vocabulary knowledge (VK)					-	0.20	0.36
6. IQ						-	0.18
7. Metacognitive knowledge (MK)							-

Note: all correlations are significant at $p < .01$

[2] The difference in -2*Loglikelihood (or Deviance) of nested models has a Chi-square distribution with a number of degrees of freedom equal to the difference in number of estimated parameters between both models.

Table 5 Multilevel Analyses with Reading Comprehension (repeatedly measured) as Dependent Variable to Establish Influence of Student-Level Variables (Language spoken at home, IQ, Gender, Vocabulary Knowledge, Metacognitive Knowledge, and Age) (N = 1129 cases/310 students)

Model:	model 5 – 0	model 5 – 1	model 5 – 2	model 5 – 3	Model 5 – 4	model 5–5	model 5–6[a]	model 5–7
Fixed part								
Intercept	***34.684(0.964)***	***34.761(0.968)***	***33.846(1.039)***	***34.007(0.978)***	***33.989(0.983)***	***34.648(0.738)***	***34.956(0.038)***	***35.000(0.659)***
Occasion (in months)	***0.156(0.036)***	***0.156(0.036)***	***1.56(0.036)***	***0.156(0.036)***	***0.156(0.036)***	***0.155(0.036)***	***0.155(0.036)***	***0.156(0.036)***
Language spoken at home (contrast only Dutch vs. Dutch half of the time)		0.610(1.332)						
Language spoken at home (contrast only Dutch vs. mostly different language)		-1.390(1.303)						
Gender (male = 0, female = 1)			**1.517(0.707)**	**1.319(0.674)**	**1.348(0.689)**	0.552(0.544)		
IQ				***0.360(0.064)***	***0.359(0.064)***	***0.240(0.052)***	***0.242(0.052)***	***0.222(0.051)***
Age					0.128(0.646)			
Vocabulary knowledge						***0.523(0.038)***	***0.527(0.038)***	***0.473(0.039)***
Metacognitive knowledge								***0.287(0.066)***
Random part (variances)								
Class	*15.851(5.879)*	*15.422(5.740)*	*15.809(5.859)*	*13.835(5.165)*	*13.865(5.174)*	*7.270(2.845)*	*7.243(2.837)*	*6.985(2.731)*
Class slope variance occasion	**0.019(0.008)**	**0.019(0.008)**	**0.019(0.008)**	**0.019(0.008)**	**0.019(0.008)**	**0.019(0.008)**	**0.019(0.008)**	**0.019(0.008)**
Class covariance slope × intercept	−0.260(0.169)	−0.260(0.168)	−0.254(0.169)	−0.217(0.157)	−0.217(0.157)	−0.258(0.128)	−0.261(0.129)	−0.276(0.129)
Student	***26.826(3.437)***	***26.725(3.429)***	***26.496(3.411)***	***23.301(3.155)***	***23.288(3.154)***	***12.673(2.309)***	***12.678(2.309)***	***11.374(2.206)***
Student slope variance occasion	**0.018(0.009)**	**0.018(0.009)**	**0.018(0.009)**	**0.017(0.009)**	**0.017(0.009)**	**0.018(0.009)**	**0.018(0.009)**	**0.018(0.009)**
Student covariance slope × intercept	0.191(0.133)	0.191(0.133)	0.172(0.133)	0.161(0.129)	0.162(0.129)	0.112(0.114)	0.118(0.114)	0.134(0.112)
Occasion (rep. measures)	***17.560(1.065)***	***17.559(1.065)***	***17.562(1.065)***	***17.580(1.066)***	***17.581(1.066)***	***17.528(1.062)***	***17.527(1.062)***	***17.508(1.061)***
Variance explained								
Class			0.003	0.125		0.476		0.036
Class slope variance occasion			-	-		-		-
Student			0.012	0.121		0.456		0.103
Student slope variance occasion			-	0.056		-		-

Table 5 (continued)

Model:	model 5 – 0	model 5 – 1	model 5 – 2	model 5 – 3	Model 5 – 4	model 5–5	model 5–6[a]	model 5–7
Occasion (rep. measures)			-	-		0.003		0.001
Deviance testing								
-2*Loglikelihood (deviance)	7179.080	7177.604	7174.540	7144.213	7144.174	6999.880	7000.898	6982.692
Difference between −2*loglikelihood			**4.54**	**30.327**	0.039	***144.294***	(-)1.018	***18.206***
Difference df			1	1	1	1	1	1
Compared to model		0	0	2	3	4	5[a]	6

[a] Model 5–6 is more parsimonious than model 5–5

Bold = p < .05; *italicized = p < .01*; ***bold and italicized = p < .001***

Multilevel analyses

As no significant random intercept variance was found at the school level (see Appendix A), models with three levels were used (occasion-, student-, and class level). Appendix A also shows that random slopes for the occasion variable both at the class and student level significantly improved model fit.

Next, the teacher-level control variables were entered in the model (see Appendix B). Neither 'teacher replacement' nor 'cancelled classes' significantly improved model fit and both were therefore omitted from further analyses. Subsequently, the student-level control variables were entered to control for differences between students at pretest. Inclusion of age, gender, and language background did not improve model fit (see Table 5[3]), whereas vocabulary knowledge, metacognitive knowledge, and IQ did. Model 5–7 (see Table 5) is therefore the model referred to as model 6 − 0 in Table 6.

To establish the effect of the treatment on growth in reading comprehension, in Table 6, we compared Model 6 − 1 (with effect on the pretest only) with Model 6 − 2, containing the interaction between occasion and treatment. This effect was not significant implying there was no effect of the treatment on growth in reading comprehension ($\Delta IGLS = 1.827$, $df = 1$, $p > .05$).

Models 6 − 3 and 6 − 4 were conditional for testing the moderating effect of vocabulary in Model 6 − 5. The moderator effect of vocabulary was tested in Model 6 − 5 as the interaction between occasion, treatment and vocabulary (see Table 6). This interaction was significant ($\Delta IGLS = 4.595$, $df = 1$, $p < .05$). Explained student-level variance on growth in reading comprehension equals 11.1% .

The interpretation of the moderating effect of vocabulary knowledge on growth in reading comprehension becomes clear when looking at regressions for different combinations of scores on the independent variables (Hox, 2010). For treatment, we used two scores (0 and 1), for occasion we used the scores 0 and 22 for the first and the last time of measurement and for vocabulary we used three scores: one standard deviation below the mean, the mean, and one standard deviation above the mean. The resulting six regression lines are presented in Fig. 1. The regression lines show that in cases where students score above average on vocabulary, there appears to be no difference between growth in reading comprehension for the control and treatment students. However, when students score below average on vocabulary knowledge, growth in reading comprehension appears to be different for control and treatment students; the control students outperform the treatment students over time. When looking at the regression lines for the treatment group only, it seems that the students scoring lower on vocabulary profit less than students scoring higher on vocabulary. In contrast, in the control group students high and low on vocabulary grow in equal measures on reading comprehension.

[3] In Table 4 the N of cases is determined by the sample N (310) times 4 (repeated measures) minus the missings on one or more of the repeated measures for reading comprehension (totalling 1129).

Discussion

This study set out to analyze the moderating effect of vocabulary knowledge on growth in reading comprehension of adolescents with a low academic achievement in a two-year intervention study. The intervention consisted of a program directed at the application of reading strategies. The control group followed lessons as usual at their schools. Our aim was to analyze whether students in the control and experimental conditions with different levels of vocabulary knowledge had different patterns of growth. Our study revealed a significant moderating effect of vocabulary knowledge, which explained an additional 11.1% of the variance. This means that vocabulary knowledge had a differential impact on students' growth in reading comprehension in the control and experimental groups. A closer look at the growth patterns in both groups showed that a lower vocabulary level resulted in less growth in reading comprehension in the experimental condition compared to the control condition, while there appeared to be no difference in growth in reading comprehension between the two conditions for students with a higher vocabulary level. In addition, visualization of the growth patterns in both groups (Fig. 1) suggests that the experimental condition had a discriminating effect for students with different levels of vocabulary, whereas the business-as-usual condition did not.

As hypothesized, students in the experimental condition were at a disadvantage compared to control students if their vocabulary knowledge was relatively low. This is a clear sign that students with a lower vocabulary in the experimental condition did not profit from the learned strategies in their attempts to make sense of the texts used for measuring reading comprehension. In contrast, it is remarkable that the control students with a lower vocabulary did show the same growth pattern as the control students with a higher vocabulary, while this was not the case for students in the experimental condition. The subgroup of students in the experimental condition with a lower vocabulary appeared to show significantly less growth than all other subgroups, not only experimental but also control. This suggest that there was an adverse effect of the treatment for their reading comprehension development. This is an indication that vocabulary knowledge is an important prerequisite for the application of reading strategies for adolescents with low academic achievement, as was hypothesized in this study.

Additionally, lower vocabulary knowledge may lead to cognitive overload (Sweller, 1988) in the case of application of newly learned reading strategies that are not already part of the students' repertoire. Studies into reading comprehension components have suggested that limited working memory capacity may play a role in students' application of reading strategies (Van Gelderen et al., 2003; Gelderen et al., 2007). This may certainly be the case for adolescents with lower vocabulary knowledge, since they will have to deal with more meaning-related problems for building a situation model (Kintsch, 1988), while at the same time they have to apply new reading strategies, which was the case in the experimental condition. Students in the experimental condition had to direct their attention to their texts in multiple ways in order to determine which strategy should be applied and how this strategy should be applied, while simultaneously keeping an eye on the meaning of words which are needed for successful application of each strategy. These processes

Table 6 Multilevel Analyses with Reading Comprehension (repeatedly measured) as Dependent Variable to Establish Influence of Treatment over Time and Interaction between Vocabulary knowledge, Occasion and Treatment, after Correcting for Control Variables (N=1129 cases/310 students)

Model	model 6−0	model 6−1	model 6−2	model 6−3	model 6−4	model 6−5
Fixed part						
Intercept	***35.000(0.659)***	***35.509(0.830)***	***34.872(0.937)***	***34.859(0.949)***	***34.873(0.947)***	***34.891(0.949)***
Occasion (in months)	***0.156(0.036)***	***0.156(0.036)***	***0.204(0.049)***	***0.205(0.049)***	***0.205(0.049)***	***0.203(0.050)***
IQ[a]	***0.222(0.051)***	***0.223(0.051)***	***0.223(0.051)***	***0.221(0.051)***	***0.223(0.051)***	***0.223(0.051)***
Vocabulary knowledge[a]	***0.473(0.039)***	***0.473(0.039)***	***0.473(0.039)***	***0.445(0.043)***	***0.507(0.057)***	***0.547(0.060)***
Metacognitive knowledge[a]	***0.287(0.066)***	***0.293(0.066)***	***0.292(0.066)***	***0.294(0.066)***	***0.289(0.066)***	***0.288(0.066)***
Treatment (treatment=1, control=0)		−0.998(0.950)	0.270(1.318)	0.272(1.335)	0.236(1.332)	0.198(1.335)
Occasion × Treatment			−0.096(0.069)	−0.096(0.069)	−0.095(0.069)	−0.092(0.071)
Occasion × Vocabulary knowledge[a]				0.004(0.003)	0.004(0.003)	−0.001(0.004)
Treatment × Vocabulary knowledge[a]					−0.118(0.072)	−0.197(0.081)
Occasion × Treatment × Vocabulary knowledge[a]						**0.011(0.005)**
Random part (variances)						
Class	***6.985(2.731)***	***7.422(2.866)***	***6.973(2.728)***	***7.205(2.800)***	***7.184(2.785)***	***7.228(2.797)***
Class slope variance occasion	**0.019(0.008)**	**0.019(0.008)**	**0.017(0.008)**	**0.017(0.008)**	**0.017(0.008)**	**0.018(0.008)**
Class covariance slope × intercept	−0.276(0.129)	−0.301(0.134)	−0.267(0.123)	−0.285(0.127)	−0.291(0.127)	−0.299(0.130)
Student	***11.374(2.206)***	***11.364(2.205)***	***11.371(2.206)***	***11.309(2.199)***	***10.935(2.171)***	***10.901(2.166)***
Student slope variance occasion	**0.018(0.009)**	**0.018(0.009)**	**0.018(0.009)**	**0.018(0.009)**	**0.018(0.009)**	**0.016(0.009)**
Student covariance slope × intercept	0.134(0.112)	0.134(0.112)	0.134(0.112)	0.140(0.111)	0.159(0.110)	0.167(0.109)
Occasion (rep. measures)	***17.508(1.061)***	***17.508(1.061)***	***17.514(1.061)***	***17.486(1.060)***	***17.491(1.060)***	***17.463(1.058)***
Variance explained						
Class						-
Class slope variance occasion						-
Student						0.003
Student slope variance occasion						0.111
Occasion (rep. measures)						0.002
Deviance testing						

Table 6 (continued)

Model	model 6−0	model 6−1	model 6−2	model 6−3	model 6−4	model 6−5
-2*loglikelihood (deviance)	6982.692	6981.709	6979.882	6977.604	6975.013	6970.418
Difference between −2*loglikelihood	18.206	0.983	1.827	1.178	2.591	**4.595**
Difference df	1	1	1	1	1	1
Compared to model		0	1	2	3	4

Bold = p < .05; *italicized = p < .01*; ***bold and italicized = p < .001;*** [a] = variable is grand mean centered

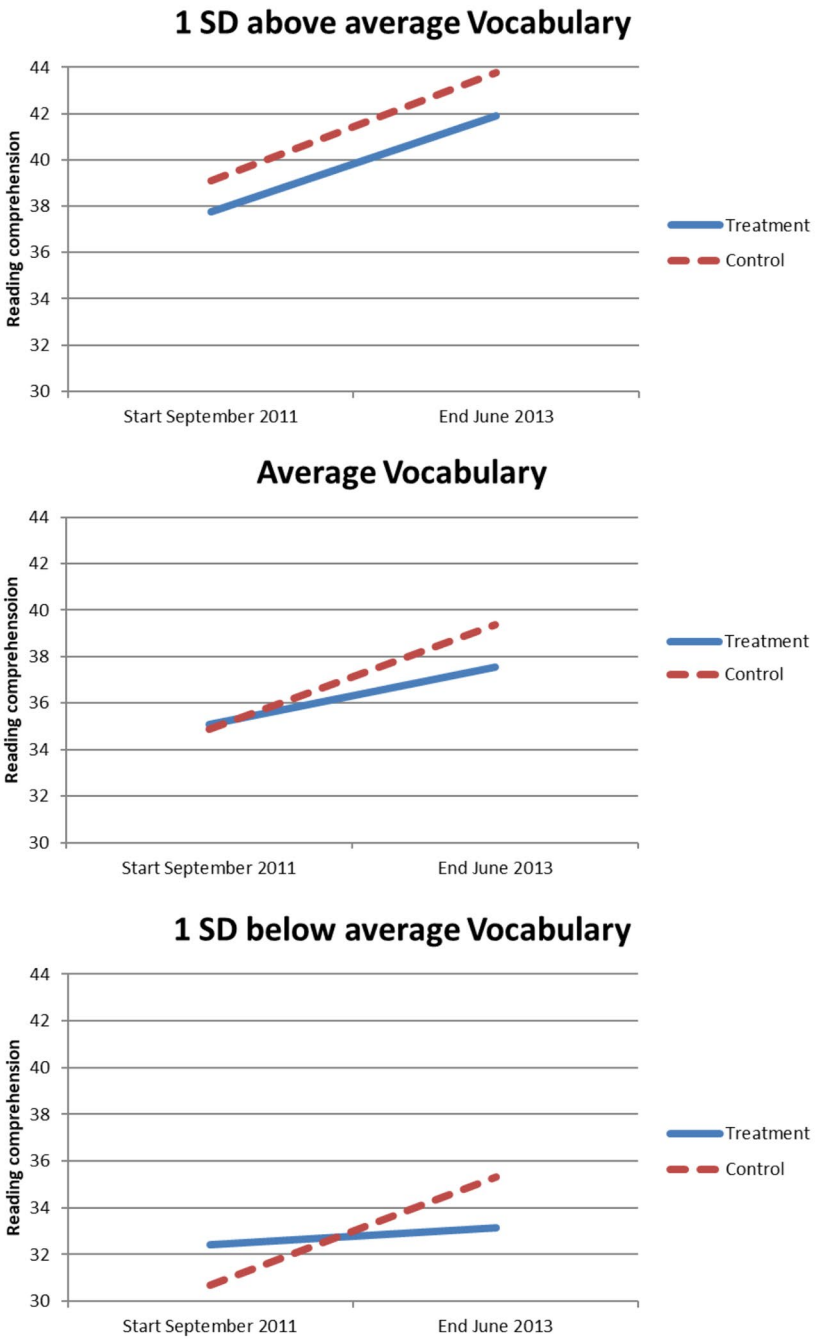

Fig. 1 Regression lines for both treatment and control, split into 1 SD above average vocabulary, average vocabulary and 1 SD below average vocabulary

may interfere with each other, resulting in a sort of "short-circuit" (Bossers, 1991; Taillefer, 1996). In that case, the search for meaning of words and sentences and the simultaneous application of (newly learned) reading strategies, such as predicting or summarizing lead to cognitive overload and failure in comprehension. This is probably a greater risk for students in the experimental condition with lower vocabulary knowledge than for students in the control condition. In the control condition, this "short-circuit" may not happen, as these students were not instructed to apply new reading strategies, but rather used strategies that were already part of their repertoire. Therefore, students in the control condition could apply reading strategies routinely, without much conscious attention.

From many studies it is known that vocabulary knowledge is a basic component of reading comprehension (Ouellette & Beers, 2010; Tilstra et al., 2009; Trapman et al., 2014; Van Gelderen et al., 2004; Gelderen et al., 2007; Van Steensel et al., 2016; Verhoeven & Van Leeuwe, 2008).

The present study adds to this knowledge by suggesting that, in the case of adolescents with low academic achievement, vocabulary knowledge is not only important in reading comprehension processes, but also plays an important role in whether students benefit from interventions directed at learning new reading strategies. The intervention tested in this study, emerged as not appropriately tailored to address the needs of adolescents with lower vocabulary knowledge, who probably need support with reading comprehension the most.

Adolescents with low academic achievement with a higher level of vocabulary were not harmed (in the sense of showing similar progress compared to the control condition), but in comparison to the control condition the students with lower vocabulary knowledge showed less progress in reading comprehension. This may also explain some of the mixed results of reading strategy interventions for adolescents with low academic achievement mentioned in the introduction (Edmonds et al., 2009; Fogarty et al., 2014; Muijselaar, 2018; Okkinga et al., 2021; Simmons et al., 2014; Vaughn et al., 2013).

Recently, researchers strive for interventions with an integrated approach in which text contents and what can be learned from them is central in discussions with students (see e.g. McKeown et al., 2009). The status of reading strategies in such interventions is that they are tools that can be used as a means, but not the goal. Attention to the meaning of words and sentences is embedded within the search of the meaning of the text as a whole. Such an approach can be beneficial for students with lower vocabulary knowledge, since it helps adding meaningful connections to their (prior) knowledge. This is corroborated by the DIME model (Ahmed et al., 2016), which suggests that vocabulary knowledge and content knowledge contribute substantially to inference-making, which in turn contributes to reading comprehension.

Suggestions for future research

Future studies into reading strategies interventions directed at adolescents with low academic achievement may profit from the following suggestions. As vocabulary knowledge is an important factor in fostering reading comprehension but may also be

important in the application of reading strategies for adolescents with low academic achievement, future studies could investigate the optimal ways to include vocabulary instruction in the context of reading strategy interventions for this group. This means that adolescents with low academic achievement should not only receive instruction in applying specific reading strategies, but also be supported in their use of vocabulary knowledge needed for successful application of reading strategies. The original approach used by Palincsar & Brown (1984) to reciprocal teaching can be used to this end. In this format, tutors support adolescents with low academic achievement in small group settings by steering group conversations directed at the application of reading strategies. The approach aims at producing a natural dialogue between the group members and the tutor, while the tutor supports all students' attempts at text comprehension. In such a small group tutoring approach, a specific vocabulary support condition allows tutors to converse intensely with the students, focusing specifically on their problems with understanding meanings of words encountered in the text. Studying the effects of such additional vocabulary support could shed light on the question whether adolescents with low academic achievement profit more from reading strategy instruction when their deficits in vocabulary are simultaneously being repaired.

Apart from vocabulary knowledge, other student characteristics can be considered as moderator variables. For example, the level of self-regulatory skills may influence how students respond to an intervention targeted at the use of reading strategies. Self-regulatory skills refer to planning, executing, and controlling behavior while performing a task (Boekaerts & Simons, 1993). In the case of reading comprehension, self-regulatory skills are important, for example, in knowing which reading strategy to apply or monitoring one's comprehension during reading. There is evidence that self-regulatory skills applied in reading tasks may determine reading comprehension of adolescents with low academic achievement (De Milliano, Van Gelderen & Sleegers,2016). Thus, the application of reading strategies (in the context of an intervention) may also be dependent on the level of self-regulatory skills of students.

In general, the fact that moderator variables such as vocabulary knowledge may play an important role in the success of reading strategy interventions makes it critical to understand how the intervention is affected by students' skills,

strengthening the need for investigating new moderator effects. This may especially be the case in whole-classroom settings, as instruction in whole-classroom settings is likely to be targeted at the average level of students, whereas instruction in small groups, as originally proposed by Palincsar & Brown (1984), is likely to be more aligned to the individual needs of the students in the small group.

Implications for educational practice

It is difficult for teachers to implement reading strategy training with reciprocal teaching for adolescents with low academic achievement in whole-classroom settings successfully (Okkinga et al., 2021). Our study confirms this difficulty, because it shows no overall advantage of the intervention in growth of reading comprehension. An important explanation for the lack of success of the whole-classroom approach is that

teachers cannot attend to multiple groups of students simultaneously and give the students the individual guidance they need. The present study deepens this explanation, as it shows that individual differences in adolescents' vocabulary knowledge influenced the effects of the experimental intervention. This suggests that other approaches are needed to facilitate reading comprehension for this group of students with lower vocabulary knowledge. In addition, the results of the control group show that it is possible to make students with lower vocabulary knowledge profit similarly to their classmates with higher vocabulary knowledge.

It is advisable for teachers involved in teaching adolescents with low academic achievement to identify students who are better served with additional vocabulary support. For this group, the focus on application of reading strategies can lead to frustration. In order to meet these students' needs, teachers can monitor their dialogues related to reading in small groups while other students are working independently. This type of classroom organization may allow teachers to give the weakest readers the attention in vocabulary support they need.

Conclusion

Results of the present study show the importance of examining moderators of the effect of reading comprehension interventions. Beyond reporting the success of interventions, we call for research that examines potential moderators for interventions, including for those that were unsuccessful. Indeed, our understanding of how to move the needle in adolescent reading comprehension can be expanded by investigating why treatments did not result in the expected effects.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11145-022-10359-2.

References

Aarnoutse, C.A.J. &; Van Leeuwe, J.F.J. (1988). Het belang van technisch lezen, woordenschat en ruimtelijke intelligentie voor begrijpend lezen [Importance of decoding, vocabulary and spatial intelligence for reading comprehension]. *Pedagogische Studiën* [*Pedagogical Studies*] *65*, 49–59

Afflerbach, P., & Cho, B. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. E. Israel, & G. G. Duffy (Eds.), *Handbook of research on reading comprehension*. New York, NY: Routledge

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, *61*, 364–373. doi: https://doi.org/10.1598/RT.61.5.1

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M., & Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7 through 12. *Contemporary Educational Psychology*, *44*, 68–82

Boekaerts, M., & Simons, P. R. J. (1993). *Leren en instructie, Psychologie van de leerling en het leerproces. [Learning and Instruction. Psychology of the student and the learning process]* Assen: Dekker & van de Vegt

Bossers, B. (1991). On thresholds, ceilings and short-circuits: The relation between L1 reading, L2 reading and L2 knowledge. *AILA Review*, *8*, 45–60

Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research*, *61*(2), 239–264. doi: https://doi.org/10.3102/00346543061002239

Dole, J. A., Nokes, J. D., & Drits, D. (2009). Cognitive strategy instruction. In S. E. Israel, & G. G. Duffy (Eds.), *Handbook of research on reading comprehension*. New York, NY: Routledge

Dutch Education Inspectorate. (2008). *Basisvaardigheden taal in het voortgezet onderwijs: Resultaten van een Inspectieonderzoek naar taalvaardigheid in de onderbouw van het vmbo en praktijkonderwijs* . Utrecht: Dutch Education Inspectorate

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A Synthesis of Reading Interventions and Effects on Reading Comprehension Outcomes for Older Struggling Readers. *Review of Educational Research*, *79*(1), 262–300. doi: https://doi.org/10.3102/0034654308325998

Field, A. P. (2009). *Discovering statistics using SPSS: and sex and drugs and rock 'n' roll*. London: Sage publications. (third edition)

Fogarty, M., Oslund, E., Simmons, D., Davis, J., Simmons, L., Anderson, L., Clemens, N., & Roberts, G. (2014). Examining the effectiveness of a multicomponent reading comprehension intervention in middle schools: A focus on treatment fidelity. *Educational Psychology Review*, *26*(3), 425–449. doi: https://doi.org/10.1007/s10648-014-9270-6

Gille, E., Loijens, C., Noijons, J., & Zwitser, R. (2010). *Resultaten PISA-2009, Praktische kennis en vaardigheden van 15-jarigen [Results PISA-2009, Practical knowledge and skills of 15-year olds]*. Arnhem: Cito

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Volume III* (pp. 403–422). New York: Erlbaum

Hazenberg, S., & Hulstijn, J. H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, *17*, 145–163. doi: https://doi.org/10.1093/applin/17.2.145

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications. Second Edition*. New York: Routledge

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441–480. doi: https://doi.org/10.1016/0010-0285(76)90015-3

Just, M. A., & Carpenter, P. A. (2004). A theory of reading: From eye fixations to comprehension. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading. Fifth edition* (pp. 1182–1218). Newark, DE: International Reading Association

Kelly, M., Moore, D. W., & Tuck, B. F. (2001). Reciprocal teaching in a regular primary school classroom. *Journal of Educational Research*, *88*, 53–61. doi: https://doi.org/10.1080/00220671.1994.9944834

Kendeou, P., Van Den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning disabilities research & practice*, *29*(1), 10–16

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. doihttps://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323. doi: https://doi.org/10.1016/0010-0285(74)90015-2

Logan, S., & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where the differences lie. *Journal of Research in Reading*, *32*, 199–214. doi: https://doi.org/10.1111/j.1467-9817.2008.01389.x

McKeown, M. G., Beck, I. L., & Blake, R. G. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading research quarterly*, *44*(3), 218–253

de Milliano, I., & Gelderen, A. (2016). Types and sequences of self-regulated reading of low-achieving adolescents in relation to reading task achievement. *Journal of Research in Reading*, *39*(2), 229–252. doi:10.1111_1467-9817.12037. SleegersP.

Ministry of Education, Culture, &, & Science (2006). *The education system in the Netherlands 2006*. The Hague: Ministry of Education, Culture, & Science/Dutch Eurydice Unit

Muijselaar, M., Swart, N., Steenbeek-Planting, E., Droop, M., Verhoeven, L., & de Jong, P. (2018). The effect of a strategy training on reading comprehension in fourth-grade students. *The Journal of Educational Research*, *111*(6), 690–703

Okkinga, M., van Gelderen, A. J. S., Van Schooten, E., van Steensel, R., & Sleegers, P. J. C. (2021). Implementation quality of principles of reciprocal teaching in whole-classroom settings: a two-year study with low-achieving adolescents. *Reading Psychology*, *42*(4), 323–363

Okkinga, M., van Steensel, R., van Gelderen, A. J., & Sleegers, P. J. (2018). Effects of reciprocal teaching on reading comprehension of low-achieving adolescents. The importance of specific teacher skills. *Journal of research in reading*, *41*(1), 20–41

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: how oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, *23*, 189–208. doi: https://doi.org/10.1007/s11145-008-9159-1

Organisation for Economic Co-operation and Development [OECD]. (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD

Organisation for Economic Co-operation and Development [OECD]. (2014). "Profile of student performance in reading". *PISA 2012 Results: What students know and can do (Volume I, Revised edition, February 2014): Student performance in mathematics, reading and science*. OECD Publishing

Palincsar, A. S., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*, 117–175. doi: https://doi.org/10.1207/s1532690xci0102_1

Palincsar, A. S., Brown, A., & Martin, S. M. (1987). Peer interaction in reading comprehension instruction. *Educational Psychologist*, *22*, 231–253. doi: https://doi.org/10.1080/00461520.1987.9653051

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of functional literacy*, *11*, 67–86

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook*. Oxford: Blackwell. doi: https://doi.org/10.1002/9780470757642.ch13

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

Raphael, T. E., George, M., Weber, C. M., & Nies, A. (2009). Approaches to teaching reading comprehension. In S. E. Israel, & G. G. Duffy (Eds.), *Handbook of research on reading comprehension*. New York, NY: Routledge

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2009). *A user's guide to MlwiN. Version 2.10*. Bristol: University of Bristol, Centre for Multilevel Modelling

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview*. San Antonio, TX: Harcourt Assessment

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research*, *64*(4), 479–530. doi: https://doi.org/10.3102/0034654306400447

Rumelhart, D. E. (2004). Toward an interactive model of reading. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading. Fifth edition* (719–747). Newark, DE: International Reading Association. doi: https://doi.org/10.1598/0872075028.41

Samuels, S. J. (2004). Toward a theory of automatic information processing in reading, revisited. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading. Fifth edition* (1127–1148). Newark, DE: International Reading Association. doi: https://doi.org/10.1598/0872075028.4

Schaffner, E., Philipp, M., & Schiefele, U. (2016). Reciprocal effects between intrinsic reading motivation and reading competence? A cross-lagged panel model for academic track and nonacademic track students. *Journal of Research in Reading*, *39*, 19–36. doi: https://doi.org/10.1111/1467-9817.12027

Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading*, *3*, 257–279. doi: https://doi.org/10.1207/s1532799xssr0303_4

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin

Simmons, D., Fogarty, M., Oslund, E. L., Simmons, L., Hairell, A., Davis, J., Anderson, L., Clemens, N., Vaughn, S., Roberts, G., Stillman, S., & Fall, A. (2014). Integrating content knowledge-building and student-regulated comprehension practices in secondary English arts classes. *Journal of Research on Educational Effectiveness*, *7*, 309–330. doi: https://doi.org/10.1080/19345747.2013.836766

Spörer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction*, *19*, 272–286. doi: https://doi.org/10.1016/j.learninstruc.2008.05.003

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, *12*, 257–285. doi:https://doi.org/10.1207/s15516709cog1202_4

Taillefer, G. F. (1996). L2 Reading ability: Further insight into the short-circuit hypothesis. *The Modern Language Journal*, *80*, 461–477. doi:https://doi.org/10.1111/j.1540-4781.1996.tb05466.x

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: components of the simple view of reading across grade levels. *Journal of research in reading*, *32*(4), 383–401. doi: https://doi.org/10.1111/j.1467-9817.2009.01401.x

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice*, *15*(10), 55–64. doi: https://doi.org/10.1207/SLDRP1501_6

Trapman, M. J. W. (2015). *Reading and writing development of low-achieving adolescents. The roles of linguistic knowledge, fluency, and metacognitive knowledge*. Amsterdam: University of Amsterdam

Trapman, M. J. W., van Gelderen, A., van Steensel, R., van Schooten, E., & Hulstijn, J. (2014). Linguistic knowledge, fluency and meta-cognitive knowledge as components of reading comprehension in adolescent low achievers: differences between monolinguals and bilinguals. *Journal of Research in Reading*, *37*, 3–21. doi: https://doi.org/10.1111/j.1467-9817.2012.01539.x

Trapman, M., van Gelderen, A., van Schooten, E., & Hulstijn, J. (2017). Reading comprehension level and development in native and language minority adolescent low achievers: Roles of linguistic and metacognitive knowledge and fluency. *Reading and Writing Quarterly*, *33*(3), 239–257. DOI: https://doi.org/10.1080/10573569.2016.1183541

Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Snellings, P., Simis, A., et al. (2003). Roles of linguistic knowledge, metacognitive knowledge and processing speed in L3, L2 and L1 reading comprehension: A structural equation modeling approach. *International Journal of Bilingualism*, *7*(1), 7–25. doi:https://doi.org/10.1177/13670069030070010201

Van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second-language reading comprehension: a componential analysis. *Journal of Educational Psychology*, *96*(1), 19–30. doi: https://doi.org/10.1037/0022-0663.96.1.19

Van Gelderen, A., Schoonen, R., Stoel, R. D., De Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in Language 1 and Language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology*, *99*, 477–491. doi: https://doi.org/10.1037/0022-0663.99.3.477

Silfhout, G., Evers-Vermeul, J., Mak, W. M., & Sanders, T. J. M. (2014). Connectives and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology*, *4*, 1036–1048. doi: https://doi.org/10.1037/a0036293

Van Steensel, R., Oostdam, R., & Van Gelderen, A. (2013). Assessing reading comprehension in adolescent low achievers: subskills identification and task specificity. *Language Testing*, *30*(1), 3–21. doi: https://doi.org/10.1177/0265532212440950

Van Steensel, R., Oostdam, R., Van Gelderen, A., & Van Schooten, E. (2016). The role of word decoding, vocabulary knowledge and meta-cognitive knowledge in monolingual and bilingual low-achieving adolescents' reading comprehension. *Journal of Research in Reading*, *39*, 312–329. doi: https://doi.org/10.1111/1467-9817.12042

Vaughn, S., Roberts, G., Klingner, J. K., Swanson, E. A., Boardman, A., Stillman-Spisak, S. J., et al. (2013). Collaborative strategic reading: Findings from experienced implementers. *Journal of Research on Educational Effectiveness*, *6*, 137–163. 10.1080/ 19345747.2012.741661

Verhoeven, L., & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology*, *22*, 407–423. doi: https://doi.org/10.1002/acp.1414

Webb, S., Massey, D., Goggans, M., & Flajole, K. (2019). Thirty-five years of the gradual release of responsibility: Scaffolding toward complex and responsive teaching. *The Reading Teacher*, *73*(1), 75–83. https://doi.org/10.1002/trtr.1799

Authors and Affiliations

M. Okkinga[1] · A. J. S. van Gelderen[1,2] · E. van Schooten[1,3] · R. van Steensel[3,4] · P. J. C. Sleegers[5]

✉ M. Okkinga
mariska.okkinga@gmail.com

A. J. S. van Gelderen
A.J.S.van.gelderen@hr.nl; avangelderen@kohnstamm.uva.nl

E. van Schooten
evanschooten@kohnstamm.uva.nl; e.j.van.schooten@hr.nl

R. van Steensel
roelvansteensel@gmail.com

P. J. C. Sleegers
petersleegers@bmc.nl

1 Rotterdam University of Applied Sciences, Museumpark 40, 3015CX Rotterdam, The Netherlands

2 Kohnstamm Institute, P.O. Box 94208, 1090 GE Amsterdam, The Netherlands

3 Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

4 Faculty of Humanities, Free University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

5 BMC Advice, P.O. Box 490, 3800AL, Amersfoort, The Netherlands