What's the difference?

Identifying characteristics
of good and poor comprehenders

Master Thesis

Bo Derksen -0941867-

Leiden University, Faculty of Social Science

What's the difference?

Analyzing characteristics of good and poor comprehenders

Master Thesis

D.G. (Bo) Derksen (0941867)

d.g.derksen@umail.leidenuniv.nl

Supervisor

A. Helder, MSc.

a.helder@fsw.leidenuniv.nl

Second Reader

Prof. P.W. van den Broek

broekpwvanden@fsw.leidenuniv.nl

Preface

With this master thesis I conclude my studies at Leiden University. All involved know that writing the final document did not come easy, and I would like to thank everyone for their full support during my graduation period.

Bo Derksen
Delft, May 2012

Contents

Abstract 7
Introduction 9
Methods 11
Participants 11
Instruments 11
Procedures 14
Results 15
Data Inspection 15
Data Analysis 15
Discussion 18
References 21

Abstract

Readers can be poor comprehenders in the face of sufficient basic reading skills. Common perception is that poor comprehenders have difficulty constructing a coherent mental representation of the text in their memory. In current study, among good (N = 11) and poor comprehenders (N = 14), was analyzed through one-way ANOVA whether good and poor comprehenders draw different inferences and if there are relations between inference making patterns for good and poor comprehenders. The analysis showed that poor comprehenders engage significantly different online processes, and show weaker working memory abilities compared to good comprehenders. This research provides further evidence that working memory abilities influence reading comprehension.

Students who have difficulty reading are more likely to lag behind in school (Vernooy, 2009). Since reading skills are addressed in at least 85% of the curriculum (Lyon, 2001), it is evident that there is a strong relation between academic success and reading proficiency. Even though the public is aware of the importance of reading, and much is done to promote reading proficiency, many students still struggle with reading. In the Netherlands 25% of the students leave primary school with a reading level two years behind the determined end level (Inspectie van het Onderwijs, 2007, 2008). Furthermore, over 30% of the students fail a national standardized reading test in the final year of primary school (Cito, 2010).

Previous decades numerous interventions have been designed to aid students' reading development. However, most of the interventions are expensive and time consuming, because they take place outside the classroom. Additionally, most interventions focus on teaching poor readers skills good readers use. However, studies have shown that strategies of poor and good readers differ (e.g., Cain & Oakhill, 2006; Nation, Clark, & Snowling, 2002; Stothard & Hulme, 1996), which means that the interventions might teach poor readers skills which do not apply for the type of strategy they use. In addition, various skills make up a good reader and it is unlikely poor readers lack all of these skills (e.g. Cain & Oakhill, 2006; Nation, Clark, & Snowling, 2002; Stothard & Hulme, 1996). These considerations make it unlikely that teaching poor readers strategies good readers use is the most effective way to intervene.

Plausibly a more effective intervention would be tailored to the characteristics of poor readers and their specific needs. In order to design these interventions a clear concept of the specific characteristics of poor readers is needed. In relation to this aspect, an important realization is that poor readers can have difficulty reading for different reasons. Some poor readers have difficulties decoding and understanding word meaning, while others possess basic reading skills but have difficulty making the right connections (e.g. linking current information to aspects mentioned in the text or linking current information to prior knowledge) and integrating this in the reading process (Rapp et al., 2007). The latter are called poor comprehenders. While for students with basic reading skill deficits causes for reading problems are more straightforward (for example, their reading difficulties are caused by poor vocabulary knowledge or decoding skills) the underlying causes for poor comprehenders' difficulties with reading remain uncertain (Cain & Oakhill, 2006). Poor

comprehenders have difficulty grasping the gist of a text in the face of age-appropriate word meaning and decoding skills (Rapp et al., 2007).

Several studies have examined characteristics of poor comprehenders (e.g., Cain, Oakhill, & Bryant, 2004; Cain, Oakhill, & Lemmon, 2004; Stothard & Hulme, 1995), since understanding which deficiencies are at the root of poor comprehension can be the base for designing effective interventions. All studies mentioned reported characteristics for heterogeneous groups of poor comprehenders, however not all poor comprehenders seem lack the same skills. For example: Nation, Clark, & Snowling (2002) found that most poor comprehenders had more difficulty recalling sentences and inserting the correct form of the irregular past tense in a cloze task than good comprehenders. Still, only three out of 23 poor comprehenders scored below-average on both tasks. In addition, Cain & Oakhill (2006) examined a group of poor comprehenders. Their results showed a heterogeneous group of poor comprehenders as well, however their study did not report any skills on which all poor comprehenders scored below the mean, indicating that there does not seem to be one clear cause for poor comprehenders' reading difficulties.

Possibly these studies could not report any homogeneous characteristics for poor comprehenders, since they focussed on specific reading skills offline at performance level. Offline means that reading performance was measured after the reading exercise was concluded. Recent studies illustrated that examining poor comprehenders' characteristics online, during reading, might enable us to develop a clearer concept of the causes for poor comprehenders' difficulties (Rapp et al., 2007). Van den Broek et al. (2006) found that task output for poor comprehenders might differ because poor comprehenders employ different cognitive processes during reading. In the study poor comprehenders performed an online reading task, and the outcomes presented two distinguishable types of responses for poor comprehenders. Van den Broek et al. (2006) reported typical errors for both subgroups, indicating that one subgroup made errors the other subgroup did not and vice versa. Several other studies reported the same similarities within subgroups of poor comprehenders (McMaster et al., 2010; Rapp et al., 2007).

Several studies also support the concept of weaknesses in cognitive inhibition as a characteristic for poor comprehenders. For example, De Beni and Palladino (2000) compared the suppression efficiency of good and poor comprehenders and poor comprehenders made more intrusion errors on a memory task. Poor comprehenders also produced more irrelevant information when asked to recall a text they read previously. Work by Carretti et al. (2005) and Cain (2006) supports the idea that poor comprehenders have problems with working

memory, and that these working memory deficits are associated with difficulties in suppressing irrelevant information. Both studies revealed inhibition deficits in memory in poor comprehenders.

The goal of present study was to examine inference making patterns for good and poor comprehenders, in order to (a) identify whether good and poor comprehenders draw different inferences (b) how these different patterns relate to vocabulary and working memory, and (c) whether poor comprehenders have more difficulty understanding a text due to weaker inhibition (e.g. poor comprehenders remember more irrelevant information, and have more difficulty filtering irrelevant from relevant information). The hypothesis is that good and poor comprehenders employ different online processes and therefore draw different inferences, and that poor comprehenders have weaker vocabulary skills and working memory and inhibition abilities compared to good comprehenders.

Methods

Participants

Participants for this study were primary school children (N = 25), both male (N = 12) and female (N = 13), from four different primary schools in the Rotterdam area in the Netherlands. The average age of participating children was 8 years 8 months. The children were selected based on their performance on a national Dutch technical reading test (Drie-Minuten-Toets (DMT)) and reading comprehension test (Begrijpend Lezen BL) (Cito, 2010). All selected children performed average or above average on DMT. Children who scored within the 40% lowest range on CITO BL were categorized as *poor comprehenders* (N = 14). Among the group of poor comprehenders were 7 male and 7 female students. Children who scored within the 40% highest range on CITO BL were categorized as *good comprehenders* (N = 11). Among the group of good comprehenders were 5 male and 6 female students. All children diagnosed with, or suspected of having any kind of reading, learning and/or attention deficit were excluded from participation.

Instruments

Estimate of cognitive capabilities. Raven's Standard Progressive Matrices (Raven SPM) (Raven, 1981) was used to measure children's nonverbal reasoning ability. There were five sets administered, with 12 items per set. Each item consisted of a target matrix with one missing part. Children were asked to select the item that best fit the matrix among six to eight

choices, with a maximum score of 60 for this task. The dependent variable was the number of correct items given in the 30 minutes children were allowed to spend on the task.

Vocabulary. The Peabody Picture Vocabulary Test (PPVT) (Schlichting, 2005), a Dutch version of the test was altered to enable group administered testing. Instead of picking out the correct answer, children were asked to circle the correct answer on a answer sheet. The test was used as an indicator of receptive vocabulary. Participants' scores were the number of correct items given in the 15 minutes. The PPVT consisted of a total of 60 items, which were increasing in difficulty. In addition, a vocabulary-matching test was administered during individual testing. This 20-item test required children to match words with meaning. Participants' scores were the number of correct items given in 3 minutes.

Reading Fluency. The children were given a Curriculum Based Measurement (CBM) task to assess oral reading fluency (Deno, 1985). In this task, children read aloud a text for one minute. The participants' scores were the number of words children read correctly minus the number of words children read incorrectly in one minute.

Working Memory. The Reading Span Task (translated and adapted from Swanson, 1992) was administered. Children were asked to remember the content and the last word of groups of sentences, read by the examiner. The number of sentences gradually increased. After reading the sentences, children responded to a content question about one of the sentences and were asked to recite the last word of each sentence. The participants' scores were the number of complete sets, and in addition, in order to further differentiate between participants, .5 points were awarded when children either recalled all last words or answered all questions correctly.

Eriksen Flanker Task. In the Eriksen Flanker Task (Flanker) (Eriksen & Eriksen, 1974), children pressed a button, as quickly and accurately as possible, corresponding with the direction of the central arrow on the screen (either left or right), ignoring any other arrows flanking the central arrow. There were a total of 70 trials. The trial types were presented in random order, with the constraint that the same condition did not occur on more than 3 successive trials. The first 20 trials were without flanking arrows, the remaining 50 trials with flanking arrows. Response time (RT) in milliseconds was the dependent variable for this task.

Differentiations were made in RT without flankers (Flanker RT), RT with consistent flankers (Flanker RTC), and RT with inconsistent flankers (Flanker RTI).

Think Aloud & Recall. For the think-aloud task the children read four texts: two narrative and two expository texts. The texts were between 163-194 words, 726-876 characters, and 14-21 sentences long. The children read the stories sentence by sentence out loud to the experimenter, and were instructed to give their comment after each sentence. After reading the texts and answering two yes/no-questions children were asked to recall the story they read as accurately as possible.

Students' responses during the think-aloud procedure were transcribed. The responses were parsed into clauses and coded. Ambiguous responses were marked and discussed with three other raters to find consensus. Each clause was categorized based on a coding scheme adapted from previous research (Kendeou & van den Broek, 2005; Linderholm & van den Broek, 2002; Pritchard, 1990). The coding scheme consisted of eleven categories: *Association (A)*, which refer to concepts from background knowledge brought to mind by the text. *Connecting Inferences* (CI) involve explaining the current sentence by connecting its meaning with preceding sentences. *Reinstatement Inferences* (RI) involve explaining the current sentence by using prior text information, not immediately preceding the sentence. *Elaborative Inferences* (EI) involve explaining the current sentence on the basis of background knowledge. When explanations were in line with the text, these were labelled valid. When explanations interfered with the text they were labelled invalid. *Predictive Inferenc*es (PI) involve anticipation of what will occur next. When predictions were in line with the text, these were labelled valid. When predictions interfered with the text they were labelled invalid. *Paraphrases* (P) and *Text Repetitions* (TR) involve putting the current sentence or part of the current sentence into own words or repeat the entire sentence. Paraphrases that did not match the current sentence were labelled *Invalid Paraphrases* (IP). *Affective Responses* (AR) indicate emotions related to the text. *Questions* (Q) involve forming a question relevant to the current sentence. Any responses that did not fall in any of the above categories were coded as *Other* (O). After labelling the clauses, ambiguous responses were marked and discussed with three other raters to find consensus.

The Recall responses were transcribed and categorized in a similar manner (Linderholm & van den Broek, 2002). The Recall coding scheme consisted of four categories: *Conservative Responses* (C) indicate a response that was literally repeated from the text, with a maximum of three altered words per sentences. *Liberal Responses* (L) indicate responses in

which sentences from the text are repeated in own words. *No Match Consistent Responses* (NMC) indicate the responses that are not mentioned in the text, but are plausible considering the text. *No Match Inconsistent Responses* (NMI) indicate responses that are not mentioned in the text, and are contradicting the text.

Procedures

The Raven SPM and PPVT were group administered consecutively. Next children were invited for one-on-one sessions during which the CBM Reading Fluency, Think Aloud, Recall, Reading Span Task and vocabulary-matching test were administered. Subsequently students finished the Erikson Flanker Task and Mental Counter Task on the computer, data of the latter has not been used in this study.

During the think-aloud procedure the children read and thought aloud about one practice narrative. The think-aloud procedure was adapted from Ericsson and Simon (1993). The experimenter performed a demonstration of the task with the first part of the practice story. During the demonstration, the experimenter gave specific examples of what to say, which were part of a rubric developed to make sure all identified cognitive processes were exemplified. These instructions were neutral, and children were prompted to tell everything that came to mind. After the demonstration children practiced thinking aloud with the final part of the practice text. Only during this part of the instruction children were prompted to elaborate, by asking following questions when children experienced difficulty thinking aloud: "What else are you thinking?" "Is there anything else you could think of when you read this sentence?".

After the instructions children were only prompted to tell what they were thinking when they forgot to think aloud after reading a card, by saying: "You forgot to think aloud.". During the procedure children read one sentence at a time and then talked to the experimenter about what they were thinking while reading the sentence in scope. After reading each story and thinking aloud, the experimenter asked children two yes/no questions about the story to check on comprehension. Directly after answering the yes/no questions children were asked to repeat the story they just read in their own words. When children indicated they had told all they remembered the experimented prompted them twice to tell more, by asking: "You've already told a lot, but do you remember more?". These one-on-one sessions were tape recorded.

Results

By performing one-way ANOVAs was analyzed whether good and poor comprehenders draw different inferences and if there are relations between inference making patterns for good and poor comprehenders.

Data-inspection

Missing data. Due to disturbed administration was decided to exclude one poor comprehender's Swanson score. For another poor comprehender the Flanker and Recall data are missing due to technical errors. Additionally, for one good comprehender Recall data is missing due to technical errors.

Descriptive statistics. On several tasks good comprehenders scored higher on tasks compared to poor comprehenders, see Table 1. Significance in difference between good and poor comprehenders is indicated with an asterisk.

Table 1

Descriptive statistics

	Good Comprehenders					Poor Comprehenders				
	N	Min	Max	Mean	SD	*N*	Min	Max	Mean	SD
# correct on Raven	11	27	41	34	4.22	14	13	40	29.36	7.25
% correct on PPVT *	11	66.67	88.33	76.96	7.45	14	48.33	78.33	66.19	8.81
% correct on vocabulary-matching	11	5	45	27.73	12.32	14	0	35	18.93	10.96
Score on Swanson Reading Span*	11	2	4	2.45	.71	13	1	3	1.71	.38
% correct on Flanker, no flankers	11	84	98	94.36	4.27	13	78	100	93.69	5.65
% correct on Flanker, consistent Flankers	11	94	100	98	2.16	13	88	100	96.15	3.60
% correct on Flanker, consistent Flankers	11	84	98	92.54	5.07	13	80	100	90.92	5.69

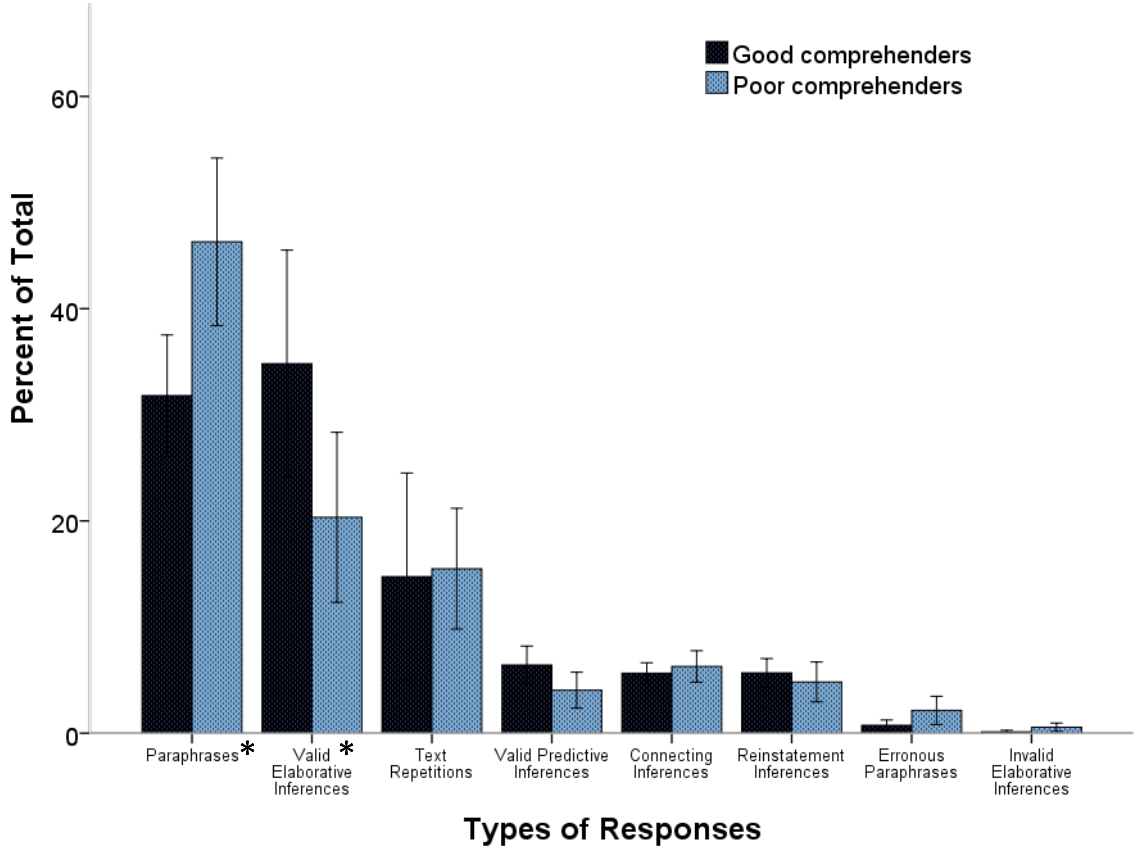

Figure 1. **Illustration of Think Aloud responses.** Percentages of types of responses during Think Aloud differentiated for good (N = 11) and poor (N = 14) comprehenders including standard error bars, significance is indicated with asterisks

Data Analysis

Think Aloud. During the Think Aloud procedure (N = 25) good comprehenders had an average of 139 responses, while poor comprehenders scored an average of 117 responses, this difference was not significant. For none of the four Think Aloud stories a significant difference was found when comparing the number of responses for good and poor comprehenders, Story 1, $F(1,22) = .366$, $p = .551$; Story 2, $F(1,22) = .903$, $p = .352$; Story 3, $F(1,23) = 2.737$, $p = .112$; Story 4, $F(1,23) = 2.014$, $p = .169$. These results indicated that there was no difference between stories and comprehension level did not influence the number of responses per individual story. When analyzing the difference in types of responses it showed that poor comprehenders made significantly more paraphrases than good comprehenders, $F(1,23) = 7.980$, $p = .010$. In addition, the difference in number of elaborate inferences differed significantly for poor or good comprehenders, $F(1,23) = 4.897$, $p = .037$.

Recall. When analyzing the difference responses for good and poor comprehenders during recall (see Figure 2) there was no quantitative response difference, $M = 57$ for poor comprehenders (N = 13) against $M = 58$ for good comprehenders (N = 10). However, the difference in recalled inconsistent responses was significant, $F(1,21) = 10.140$, $p = .004$, which indicates that comprehension level influences the number of inconsistent responses during recall. The difference in recalled conservative responses was also significant, $F(1,21) = 4.446$, $p = .047$, indicating comprehension level influences the number of conservative responses during recall.

Working memory. When analyzing the difference in responses per comprehension level significant effects were determined, $F(1,22) = 10.788$, $p = .003$. Indicating comprehension level significantly influences Reading Span scores.

Vocabulary. The vocabulary-matching test indicated no significant difference between the scores of good and poor comprehenders, $F(1,23) = 3.565$, $p = .072$. While the Peabody Picture Vocabulary Test did indicate a significant difference between the scores of good and poor comprehenders, $F(1,23) = 10.535$, $p = .004$.

Erikson Flanker Task. Differences between good and poor comprehenders in response times when arrows appeared on screen were not significant, $F(1,22) = .189$, $p = .668$. Nor were the differences for response times with consistent flankers, $F(1,22) = .253$, $p = .620$. The difference in response times for inconsistent flankers was not significant either, $F(1,22) = 1.393$, $p = .251$, indicating that comprehension level did not influence measured response times for the Flanker Task.

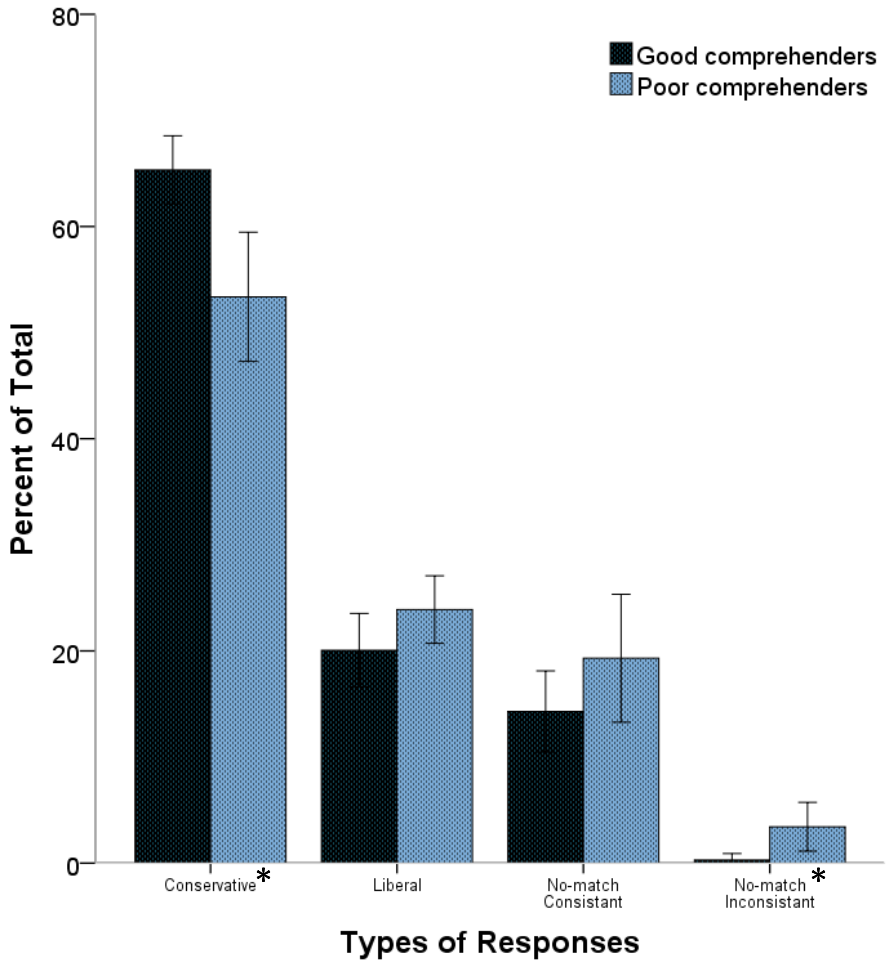

Figure 2. **Illustration of Recall responses**. Percentages of types of responses during Recall differentiated for good (N = 10) and poor (N = 13) comprehenders including error bars, significance is indicated with asterisks

Discussion

The goal of the study was to examine inference making patterns and examine patterns for good and poor comprehenders, in order to (a) identify whether good and poor comprehenders draw different inferences (b) how these different patterns relate to vocabulary and working memory, and (c) whether poor comprehenders have more difficulty understanding a text due to weaker inhibition (e.g. poor comprehenders remember more irrelevant information, have more difficulty filtering irrelevant from relevant information). Results of this present study showed that poor and good comprehenders draw different inferences. Good comprehenders make more elaborate inferences than poor comprehenders, which could indicate that good comprehenders make more use of their background knowledge

in order to comprehend the text. In addition, poor comprehenders made more paraphrases. This indicates that poor comprehenders summarize the text during reading, which is a form of comprehension monitoring.

In addition, when participants were asked to recall the texts they previously read, good and poor comprehenders responded differently. Poor comprehenders had more difficulty accurately recalling the texts. On average poor comprehenders scored a higher number of no-match inconsistencies. An example of a no-match inconsistency is when a child recalls that a sweater was blue, while it was actually red and the red color was somehow significant for the gist of the story. The higher number of no-match inconsistencies for poor comprehenders indicates that poor comprehenders recalled fragments which were not represented in the text, and actually contradicted the text. Therefore, it could be stated that poor comprehenders showed to have difficulty accurately grasping the gist of the text. These findings support the common perception that reading comprehension is influenced by the ability to construct a coherent mental representation of the text in readers' memory (Graesser, McNamara, & Louwerse, 2003; Kintsch, 1998). Good comprehenders showed to have less difficulty constructing coherent mental representations of the text, as they were able to recall a significantly higher number of conservative nodes. Indicating that good comprehenders could recall more accurate phrases from the text they just read than poor comprehenders.

When analyzing vocabulary results for good and poor comprehenders it showed not significant difference in responses for good and poor comprehenders. However, poor comprehenders scored significantly lower on the Peabody Picture Vocabulary Test. Indicating a difference in vocabulary knowledge for good and poor comprehenders. These findings suggest vocabulary knowledge can influence comprehension. However, other research has shown vocabulary does not exclusively clarify readers' poorer comprehension (Borella, Carretti & Pelegrina 2010; Locacsio et al., 2010).

Working memory scores illustrated a significant difference in good and poor comprehenders. Poor comprehenders scored significantly poorer on the Reading Span task and recalled fewer words compared to good comprehenders. In addition, the overall response times on the Flanker task indicated no difference in response times for good and poor comprehenders. These findings contradict studies which have indicated that poor comprehenders have difficulty suppressing irrelevant information (Borella, Carretti & Pelegrina 2010) and therefore have more difficulty with text comprehension.

On a whole present finding support earlier findings that there is a relation between comprehension and inference making patterns, which do not appear to stem solely from vocabulary knowledge. Current findings also suggest poor comprehenders have underlying inefficient cognitive inhibition. Plausibly, poor comprehenders' weaknesses in regulating the contents of working memory cause difficulties with reading comprehension, as supported by Pimperton and Nation (2010).

However interesting current findings present study was performed with only a small number of participants. For future studies it seems profitable to analyze whether differences in characteristics for good and poor comprehenders are still apparent when comparing and testing a larger number of participants.

References

Borella, E., Carretti B., &, Pelegrina S. (2010) The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*, *4*, 541–552.

Cain, K. (2006). Children's reading comprehension: The role of working memory in normal and impaired development. *Working memory and education*, *3*, 61–91.

Cain, K. & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology, 76*, 683-696.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31-42.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: the influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*, 671-681.

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A metaanalysis. *Learning and Individual Differences*, *19*, 246–251.

De Beni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks – Are they related to reading comprehension ability? *Learning and Individual Differences*, *12*, 131–143.

Deno, S. L. (1985). Curriculum-based measurement: the emerging alternative. *Exceptional Children*, *52*, 219-232.

Ericsson K. A. & Simon H. A. (1993). Protocol Analysis: Verbal Reports as Data MIT Press, Cambridge, MA.

Eriksen B. & Eriksen, C. (1974). Effect of noise letters upon the identification of target letters in a non search tasks. *Perception and Psychophysics*, *16*, 143-149.

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text. In A. P. Sweet, & C. E. Snow (Eds.), Rethinking reading comprehension, 82-98.

Kendeou, P. & van den Broek, P. (2005). The role of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, *2*, 235–245.

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

Locascio, G., Mahone, M., Eason, S. H., Cutting, L. E. (2010) Executive dysfunction among children with reading comprehension deficits. *Journal of Learning Disabilities*, *5*, 441–454.

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology*, *94*, 778-784.

Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., Schulte, A., & Olson, R. (2001). *Rethinking learning disabilities*. Washington, D.C.: Hudson.

Inspectie van het onderwijs (2007). De staat van het Onderwijs. Onderwijsverslag 2005/2006. Utrecht: inspectie van het onderwijs.

Inspectie van het onderwijs (2008). Basisvaardigheden taal in het voortgezet onderwijs. *Resultaten van een inspectieonderzoek naar taalvaardigheid in de onderbouw van het vmbo en praktijkonderwijs.* Utrecht: inspectie van het onderwijs.

McMaster, K., van den Broek, P., White, M. J., Rapp, D. N., Kendeou, P., Espin, C., Carlson, S., & Bohn-Gettler, C. M. (2010). Making the Right Connections: What Types of Questions Promote Struggling Readers' Comprehension? Manuscript submitted for publication.

Nation, K., Clarke, P., Snowling, M. J. (2002). General cognitive ability in children with reading comprehension difficulties. British Journal of Educational Psychology, *72*, 549-560.

Pimperton, H., & Nation, K. (2010). Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. *Journal of Memory and Language*, *62*, 380-391.

Pritchard, R. (1990). The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly*, *4*, 273–295.

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*, 289-312.

Raven, J. (1981). Manual for Raven's Progressive Matrices and Vocabulary Scales. Research Supplement No.1: The 1979 British Standardisation of the Standard Progressive Matrices and Mill Hill Vocabulary Scales, Together With Comparative Data From Earlier Studies in the UK, US, Canada, Germany and Ireland. San Antonio, TX: Harcourt Assessment

Schlichting, L. (2005). Peabody Picture Vocabulary Test-III-NL. Amsterdam: Harcourt Test Publishers.

Stichting Cito Instituut voor Toetsontwikkeling Arnhem (2010). Wetenschappelijke verantwoording. DMT en AVI.

Stothard, S. E., & Hulme, C. (1995). A comparison of phonological skills in children with reading comprehension difficulties and children with decoding difficulties. *Journal of Child Psychology and Psychiatry*, *36*, 399–408.

Swanson, H. L. (1992). Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology*, *84*, 473-488.

Van den Broek, P., McMaster, K., Rapp, D.N., Kendeou, P., Espin, C., & Deno, S. (2006). Connecting cognitive science and educational practice to improve reading comprehension. Paper presented at the Institute of Education Sciences Research Conference. Washington, D.C.

Vernooy, K. (2009). Leesonderwijs basisonderwijs behoeft dringend verbetering. In: Christen Democratische Verkenningen special 'De nieuwe schoolstrijd".