

Differences in the Effects of Summarizing Skills Training by 4th Grade Students

Tina Pirc, Sonja Pečjak

Department of Psychology, Faculty of Arts, University of Ljubljana, Slovenia

The study presents the effects of a 6-month intervention program for training students in using summarizing skills by informational texts. A total of 114 4th grade students and 5 of their teachers, who implemented the program, participated in the experimental group (EG), 76 students were in the control group (CG). We examined the students' skill of summarizing with a pre-, post- and a follow-up test. All students were divided into four groups according to their general reading competency (GRC) at pre-test: struggling, at-risk, average and good readers. We found: 1) important progress in summarizing by all groups of readers (EG and CG), with a more extensive progress made by readers in the EG; 2) stable proficiency differences between students in summarizing in general and in the elements of the summary (main ideas and coherence of text). The implications for further research and practice are discussed.

Key words: summarizing, differences between students, informational texts, intervention

Introduction

During the first years of schooling, students acquire basic reading skills (reading fluency and vocabulary), which enable good reading comprehension. In upper primary school, they start encountering informational texts in their textbooks, from which they are supposed to learn. According to the learning standards for the 4th and the 5th grade in many countries, students should be able to find important information in texts and summarize it (e.g., Common Core State Standards, 2014; Curriculum for the Slovene language, 2011). With regard to these standards, summarizing is one of the most effective strategies students can use.

Correspondence concerning this article should be addressed to Tina Pirc, University of Ljubljana, Faculty of Arts, Ljubljana, Slovenia, Department of Psychology, Aškerčeva 2, 1000 Ljubljana, SI-Slovenia. E-mail: tina.pirc@ff.uni-lj.si

Received April 17, 2018

Summarizing and Differences Between Students

Summarizing strategy is a learning strategy, by which students find important information in a text and combine it into a short, coherent text – summary. To be able to do this, students have to analyze each of the sentences/paragraphs, search for important words, leave out the unimportant or specific information and then gather the important information into a whole that makes sense (Westby, Culatta, Lawrence, & Hall-Kenyon, 2010). The research findings by Kintsch (1974), and Kintsch and van Dijk (1978) offer a good foundation for understanding the summarizing process. Authors propose that each text comprises information on three levels: the most important statements are on the first level (i.e., macrostructures), statements with more details are on the second level and statements with the most details are on the third level (i.e., microstructures). They describe three processes, which are involved

in summarizing – deletion (unimportant and redundant information), generalization (using more general concepts) and integration into a joint general statement.

Despite the fact that students often have to use summarizing in their schoolwork, this is a very demanding strategy for younger students (aged 9 to 10 years). There are large (inter)individual differences in students' ability to summarize. Most of them use copy-delete strategy when making a summary (Brown, Day, & Jones, 1983; Reading Quest Organization, 2017) – they read sentence by sentence and decide whether to include each into a summary or not. If they decide to include a sentence, they almost literally copy it from the text.

Summarizing and Reading Abilities

The most important factor of making a quality summary is *reading comprehension*. In younger students, this is a multiplicative function between decoding and linguistic comprehension, which involves lexical information and deriving text representation from it. This points out “that neither decoding nor linguistic comprehension is sufficient by itself, but that good reading comprehension requires both skills” (Schwanenflugel & Flanagan Krapp, 2016, p. 168). Therefore, good reading comprehension as a prerequisite for summarizing is facilitated by automatized decoding skill and well developed vocabulary. Similarly, Cromley and Azevedo (2007) describe reading comprehension as a function of prior knowledge, fluency, vocabulary, strategies and inferences. The authors propose that reading fluency and prior knowledge influence reading comprehension directly and indirectly (through vocabulary) and that reading strategies support the processes of inferring, which enable comprehension.

On account of the proposed interactive impact of different factors on reading comprehen-

sion as the base for learning how to summarize, we combined reading fluency, vocabulary and general reading comprehension into a composite variable of GRC (see Current study and Instruments).

Differences in Reading Comprehension and Summarizing

The basic variable that differentiates students in their reading comprehension is their mastery of automatized *decoding*, which is demonstrated through reading fluency (McKenna & Stahl, 2003; Nunes, Bryant, & Barros, 2012). Students, who use most of their working memory attention for decoding, have a poorer understanding of the material they read (Hintze, Mathews, Williams, & Tobin, 2002). Research shows that reading fluency is more connected with reading comprehension of 3rd and 4th grade students, whereas 5th-graders are already able to use the context to help them understand the reading material (Saarnio, Oka, & Paris, 1990).

Reading vocabulary also affects reading comprehension in upper primary school. It has a moderate direct effect on 4th (Quellette, 2006) and 5th (Pečjak, Podlesek, & Pirc, 2009) grade students' reading comprehension. In addition, by facilitating the process of decoding (Yap, Balota, Sibley, & Ratcliff, 2012) and increasing reading fluency (Nouwens, Groen, & Verhoeven, 2015), the effect is also indirect. Larger vocabulary enables readers to decipher the meaning of individual paragraphs or sentences faster (Nation, 2004).

However, there are differences between students in reading variables, which represent the foundation for summarizing, e.g., PIRLS, 2016 (Mullis, Martin, Foy, & Hooper, 2017). From primary to upper primary school, these inter-individual differences are demonstrated in three patterns (Pfost, Hattie, Doerfler, & Artelt, 2014): 1) inter-individual differences increase during

schooling – i.e. Matthew effect (Stanovich, 2000); 2) differences decrease – students with lower reading achievement compensate for certain deficits in reading; 3) reading achievement of all students increases, yet differences between students at the starting point remain in higher grades.

In longitudinal research (Aarnoutse, Van Leeuwe, Voeten, & Oud, 2001; Baumert, Nagy, & Lehmann, 2012), authors found high stability of inter-individual differences in reading between students, which lead to numerous intervention programs with the intention of improving fundamental reading skills (decoding, vocabulary) as well as programs to improve comprehension (Schwanenflugel & Flanagan Knapp, 2016). In designing our intervention program, we wanted to address the differences in students' fundamental reading skills, which were often ignored in traditional reading research (Paris, 2005).

Current Study

Summarizing proved to improve students' reading achievement significantly in programs, where it was used in a combination with other strategies (e.g., Reciprocal teaching; Palinscar & Brown, 1984); a part of cognitive and/or metacognitive strategies (e.g., McKown & Barnett, 2007); motivational strategies (e.g., CORI program; Guthrie et al., 2004). However, it is difficult to define the "pure" contribution of summarizing to final reading achievement due to simultaneous training of other strategies in the program. By focusing only on training students in summarizing as an independent strategy, we were able to do so.

In line with Suggate's (2016) recommendation »that reading interventions generally benefit all readers, although research is needed investigating effects at a more long-term follow-up to test whether and how different readers respond to reading intervention« (p. 78), we used a fol-

low-up test. Intervention studies usually report about the following groups of students: 1) normal readers, 2) at-risk readers (reading below the 50th percentile), 3) low-performing readers (reading below the 25th percentile and 4) disabled readers (reading below 10th percentile or those who have been diagnosed as having reading – IQ discrepancy of one standard deviation) (Suggate, 2016). In our study, we assessed students' reading decoding, vocabulary and comprehension before the intervention. As researchers suggest (Paris, 2005; Riddle Buly & Valencia, 2002), we combined these variables into a composite variable of GRC and divided students into quartile groups, according to their GRC: 1) Q₁ group with students' GRC below 25th percentile ("struggling readers"); 2) Q₂ group with GRC between 25th and 50th percentile ("at risk readers"); 3) Q₃ group with GRC between 50th and 75th percentile ("average readers") and 4) Q₄ group with GRC above 75th percentile ("good readers"). We were interested in the progress of individual groups of students after the intervention program, in their developmental pattern, and in the long-term effect of the intervention. Our study was based on the following assumptions:

First, all groups of students in the EG and in the CG will make progress in summarizing and will be able to create summaries with better quality at the end of the school year, compared to the time before the intervention. We expected more extensive progress of students in the EG after the program, with larger long-term training effects and stable proficiency differences between all groups of readers in summarizing. Namely, we designed the program to be helpful to all students and the teachers trained the whole classroom in summarizing.

Second, we assumed that students with better GRC (average and good readers) in the EG would make a more extensive progress in the program – we predicted the Matthew effect. There were two reasons for this assumption:

1) summarizing is a demanding thinking strategy, which predisposes an individual to gain the ability to separate important from unimportant information and to organize important information into sentences coherently and 2) GRC proved to be an important predictor of summarizing achievement (Pečjak & Pirc, 2018).

Method

Participants

In a convenient sample (teachers were directly invited to cooperate voluntarily), 190 4th grade students from eight Slovenian public schools participated, 114 in the EG (60%) and 76 (40%) in the CG. In the EG, there were 50.9% boys and 49.1% girls and 5 of their teachers; in the CG there were 51.3% boys and 48.7% girls. The average age of students at the beginning of the intervention in both groups was the same (9.27 years; $SD = 0.31$). All students had Slovene nationality, there were no students with special needs or language deficiencies involved, families' SES was not gathered.

Students in the EG and in the CG were divided into four groups according to significant differences in their GRC: struggling, at-risk, normal and good readers. In EG $F(3) = 318.05$, $p < .001$, $\eta^2 = .897$ and in CG $F(3) = 175.064$, $p < .001$, $\eta^2 = .879$.

Instruments

We used the Reading test (Pečjak, Potočnik, & Podlessek, 2011) and the Vocabulary test (Herschel, 1963) for evaluating the students' GRC. The Reading test comprises two subtests: Reading fluency and Reading comprehension. In Reading fluency (25 items, maximum score 25 points; Cronbach's $\alpha = .92$), students fill in the missing words by choosing from the four provided words to complete the sentence meaningfully. In Reading comprehen-

sion, students read five short texts and answer 20 multiple-choice questions (maximum 20 points, $\alpha = .85$). Test of Reading (Level 3 – Elementary Form) from the Herschel's Vocabulary test (1963) was translated and adapted into the Slovenian language (Toličič & Zorman, 1977). Students answer 20 questions by selecting the right word from a choice of five words (maximum 20 points, $\alpha = .88$). We combined the results from both tests into a composite variable of GRC, representing the sum of all possible scores from both instruments with a maximum of 75 points.

We assessed summarizing three times: before the intervention (pre-test – Summarizing_1), after the six month intervention (post-test – Summarizing_2) and three months after the intervention finished (follow-up test – Summarizing_3). In the pre-test and in the follow-up test we used three short informational text excerpts from science and social textbooks (from 99 to 120 words). After reading each text, students had to write a summary. We evaluated the summaries according to the adapted version of Friend's criteria (2001): 1) amount of important information in the summary (each text comprised three semantic units, representing important information in the text; the total score for all three texts was 9 points); 2) coherence of the summary (students connected the sentences in a meaningful way or not) – students did not receive any points if the summary was incoherent; 0.5 point for a partly coherent summary and 1 point for a coherent summary (the total score was 3 points); 3) title of the text (0 points for inappropriate; 0.5 points for partly appropriate and 1 point for appropriate title; the total score was 3 points). For Summarizing_1 and Summarizing_3, the total score from all three texts was 15 points. In Summarizing_2, we used one longer text with 237 words. Our aim was to determine whether the students would be able to use the learned summarizing strategy in a different context as

well – in a longer text. We evaluated these summaries by the same criteria – maximum 9 points for important ideas, 1 point maximum for coherence and 1 point maximum for the title; maximum 11 points total.

Two independent reviewers assessed the students' summaries. If their scores differed, they had to reach consensus – they discussed the differences in a meeting until a score they agreed upon was accepted. Internal consistency of the reviewers before the meeting was .86 for Summarizing_1, .87 for Summarizing_2 and .88 for Summarizing_3.

Description of the Intervention Program

The six-month intervention program was developed by the authors, based on previously designed similar programs (e.g., Guthrie et al., 2004; McKown & Barnett, 2007). It had two parts. The first part consisted of a two-month intensive training, including 14 sessions (14 texts), which took place two times a week for 30 minutes. Teachers received manuals, in which the schedule of contents was exactly described. Texts prepared in advance for the first two months of training were also included. Students were supplied with workbooks including short informational texts (50 – 170 words). The purpose of this part of the training was to teach the students how to use summarizing gradually. Therefore, teachers trained their students in: 1) finding important information in the texts; 2) marking the main ideas; 3) meaningfully connecting the important ideas into 1–2 sentences (summary). Teachers explicitly modeled what had to be done and guided each student in his/her work with regular feedback about the appropriateness of his/her summary.

The second part was less intensive and lasted for four months. In this part, students used their newly developed summarizing skills by working with texts from their textbooks

(20 texts). The main goal was that students consolidate the skills of summarizing by using them in texts, which they encounter in their regular lessons. Hence, teachers used longer texts (2–3 paragraphs) from required science or social studies textbooks, referring to the subjects they were involved with at the time. These lessons took place 1–2 times a week and students processed the texts the same way as they did during the first two months of the intensive training. However, in this phase of the program, teachers guided students mostly by frontal feedback and only occasionally supervised independent practice of individual students or pairs of students. For each lesson of both parts of the training, we defined the goals and didactic methods. Both parts took place during regular school hours.

Teachers, who taught the EG of students, attended a one-day training, in which they were thoroughly acquainted with the content and the course of the implementation of the intervention program. They also took part in a workshop, in which we simulated the training of using summarizing strategy by the students. The purpose of the training was to achieve the most standardized implementation of our program as possible.

The authors met the teachers after the first part of the intervention was finished and again at the end of the program. In these meetings, teachers provided feedback about the course of the implementation of the program.

Data Collection

After the schools agreed to take part in our study, we gathered the parents' written consents for their children to participate. In each classroom, data collection took place during regular school hours three times: in November 2016 (pre-test), in June 2017 (post-test) and in September 2017 (follow-up test).

Results

GRC in the EG and in the CG before Intervention

First, we examined, whether individual groups of readers in the EG and in the CG were comparable in their results of GRC before the intervention (Table 1). Distribution of GRC was normal in the EG ($M=35.18$; $SD=12.59$; $min=11.00$; $max=62.00$; $Skewness=0.25$; $Kurtosis=-0.59$) as well as in the CG ($M=38.42$; $SD=11.38$; $min=10.00$; $max=60.00$; $Skewness=-0.30$; $Kurtosis=-0.48$). There were no significant differences between the EG and the CG of students in GRC before the intervention $t(188)=-1.80$; $p=0.073$, the effect size was small ($d=.25$).

The results show no significant differences in GRC in any of the groups of readers between EG and CG before the intervention. The effect sizes were small (Richardson, 2011), which indicates that all groups of readers in the EG and in the CG were comparable in reading competency at pre-test.

EG and CG Students' Progress in Summarizing

First, we evaluated students' achievement in summarizing at pre-test (Summ_1) and at the

follow-up test (Summ_3) in the EG and in the CG, because the results were directly comparable due to the use of similar instruments (Table 2). Then we examined the progress different groups of readers made in both groups from pre-test to the follow-up test (Table 3).

It is evident from Table 2 that there were no significant differences in summarizing in any of the groups of readers in the EG and in the CG before the intervention (at pre-test, Summ_1). There were also no significant differences between groups of struggling and at-risk readers in the EG and in the CG at the follow-up test (Summ_3). However, with 7.4% risk rate in the group of average readers and with 8.1% risk rate in the group of good readers, the results show significantly better achievement of these groups of readers in the EG.

Table 3 shows that struggling readers in the EG and in the CG made significant progress from pre- to the follow-up test. However, there were no significant differences between both groups in summarizing achievement at any measurement time (Table 2), with small effect sizes between the groups (Richardson, 2011). In the group of at-risk readers, there were no significant differences in summarizing before the intervention program, but at-risk readers in the EG made significant progress compared to the same group in the CG, with regard to the results in summarizing before the intervention.

Table 1 *Descriptive statistics and ANOVAs for GRC for each group of readers with regard to EG and CG*

Groups of readers		<i>N</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Q1 - struggling	EG	29	19.55	4.05	0.547	1	.464	.013
	CG	13	20.61	4.86				
Q2 - at-risk	EG	28	30.68	2.33	0.155	1	.696	.067
	CG	13	31.00	2.64				
Q3 - average	EG	30	39.03	2.89	0.169	1	.683	.003
	CG	22	38.68	3.24				
Q4 - good	EG	27	52.37	6.20	2.745	1	.103	.049
	CG	28	49.93	4.65				

Note. η^2 – effect sizes

There were no significant differences between the groups of average and good readers in the EG and in the CG before the intervention (Table 2). However, at the follow-up test, average readers in the EG and good readers in the CG made significant progress.

In the next step, we examined students' achievement in summarizing a longer text at the

end of the intervention program (post-test, Summ_2). Our goal was to investigate if students were able to make a transfer of the learned summarizing skill and use it in a new situation (Table 4).

It is apparent from Table 4, that the better the readers (in the EG as well as in the CG) the higher their achievement in summarizing a longer text.

Table 2 *Descriptive statistics and MANOVAs for Summarizing 1 and 3 with regard to different groups of readers*

Groups of readers			<i>N</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	partial η^2
Q1 - struggling	Summ_1	EG	27	3.83	2.92	0.084	1	.774	.002
		CG	10	4.15	3.04				
	Summ_3	EG	27	6.57	2.65	0.163	1	.689	.005
		CG	10	6.15	3.32				
Q2 - at-risk	Summ_1	EG	23	6.43	2.36	0.211	1	.649	.006
		CG	12	6.04	2.47				
	Summ_3	EG	23	8.06	2.00	0.596	1	.445	.018
		CG	12	7.42	2.94				
Q3 - average	Summ_1	EG	25	7.00	3.00	1.173	1	.285	.028
		CG	18	7.97	2.76				
	Summ_3	EG	25	9.00	2.17	3.356	1	.074	.076
		CG	18	7.58	2.91				
Q4 - good	Summ_1	EG	23	9.19	2.81	2.459	1	.124	.050
		CG	26	7.85	3.16				
	Summ_3	EG	23	10.21	1.83	3.185	1	.081	.063
		CG	26	9.10	2.47				

Note. partial η^2 – effect sizes.

Table 3 *Differences between pre- and follow-up test of summarizing in groups of readers from EG and CG*

Groups of readers	Summ3_1		
	<i>t(df)</i>	<i>p</i>	
Q1 - struggling	EG	4.256(26)	.000***
	CG	2.469(9)	.036*
Q2 – at risk	EG	3.347(25)	.003**
	CG	1.658(11)	.125
Q3 – average	EG	3.103(24)	.005**
	CG	-0.483(19)	.635
Q4 - good	EG	1.403(22)	.174
	CG	2.627(25)	.014*

Note. $p < .05^*$; $p < .01^{**}$; $p < .001^{***}$

The results also show that students from all groups of readers in the EG made better summaries than their peers from these groups in the CG. The summarizing achievement of at-risk, average and good readers in the EG was significantly higher than the achievement of students from these groups in the CG and the effect sizes were moderate (Richardson, 2011).

Progress of Groups of Readers in the EG with Regard to Summarizing Elements

In general, we found larger progress in summarizing achievement in the quartile groups of readers in the EG. Thus, we were interested in defining how the key elements of a good sum-

Table 4 Descriptive statistics and ANOVAs for Summarizing 2 with regard to different groups of readers

Groups of readers		<i>N</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	partial η^2
Q1 - struggling	EG	29	4.34	2.15	0.079	1	.780	.002
	CG	13	4.12	3.00				
Q2 - at-risk	EG	25	5.46	2.68	5.195	1	.029*	.126
	CG	13	3.46	2.31				
Q3 - average	EG	29	6.45	2.00	6.916	1	.012*	.128
	CG	20	4.70	2.66				
Q4 - good	EG	27	8.13	2.01	16.711	1	.000***	.243
	CG	27	6.09	1.63				

Note. $p < .05^*$; $p < .01^{**}$; $p < .001^{***}$; partial η^2 – effect sizes.

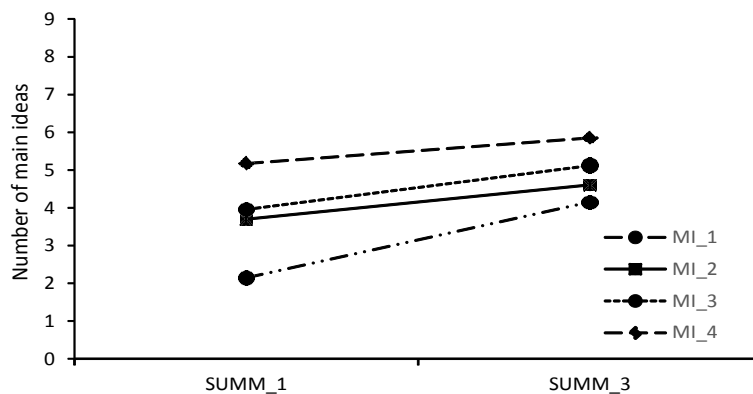


Figure 1 The average number of main ideas in summaries for individual groups of readers. MI_1(2, 3, 4) – number of main ideas in Q1(Q2, Q3, Q4).

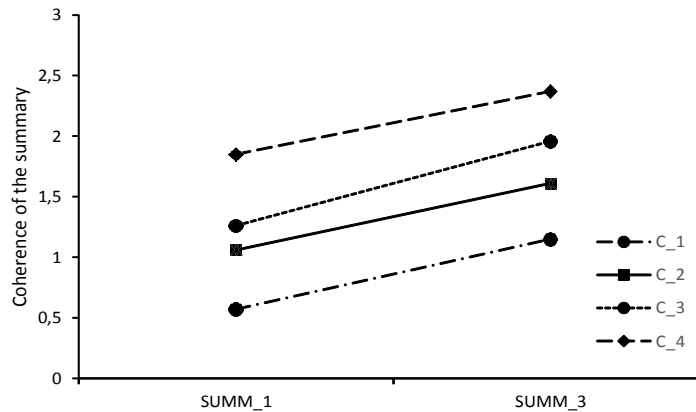


Figure 2 The coherence in summaries for individual groups of readers. C_1(2, 3, 4) – coherence of the summary in Q1 (Q2, Q3, Q4).

mary – main ideas and coherence of the text influenced this progress. In Figures 1 and 2, we present the progress of students in different quartile groups of readers in these elements from pre- (Summ_1) to the follow-up test (Summ_3).

Figure 1 shows large and significant differences between groups of students in their ability to find main ideas before the intervention $F(3) = 12.478; p < .001$; partial $\eta^2 = .263$ and moderately significant differences at the follow-up test $F(3) = 6.773; p < .001$; partial $\eta^2 = .166$. Students from Q_1 , Q_2 and Q_3 made significant progress in identifying main ideas $Q_1: t(26) = 4.447; p < .001$; $Q_2: t(25) = 3.217; p < .01$; $Q_3: t(25) = 2.926; p < .01$, whereas students in Q_4 group made progress, but it was not significant $t(22) = 1.338; p = .194$.

Figure 2 shows large and important differences between groups of readers in their ability to create a coherent summary before the intervention $F(3) = 11.896; p < .001$; partial $\eta^2 = .254$ and at the follow-up test $F(3) = 14.578; p < .001$; partial $\eta^2 = .300$. Moreover, all groups of stu-

dents made significant progress from pre- to the follow-up test, with the achievement of all students being significantly better than before the intervention $Q_1: t(26) = 3.217; p < .01$; $Q_2: t(25) = 2.872; p < .01$; $Q_3: t(24) = 3.422; p < .01$; $Q_4: t(22) = 2.709; p < .05$.

Discussion

The aim of our study was to investigate the effects of an experimental program for training 4th grade primary school students in summarizing. The program was designed on the basis of the following rationales: first, 4th grade students are faced with a transition from a period when reading is primarily aimed at developing a reading ability to a period when reading becomes a learning tool. Second, teachers' expectations that the students will be able to learn from texts by themselves increase. This means that the students have to master strategies for good comprehension. One of them is the summarizing strategy, which is an important predictor of

the students' reading achievement (Callan, Marchant, Finch, & Flegge, 2017; Hattie & Donoghue, 2016; Pečjak & Pirc, 2018).

Because of the large inter-individual differences between students in GRC, it is not reasonable to consider them as a unified group, but rather to address individual subgroups of students (see Afflerbach, 2016). Therefore, we divided students in the EG and in the CG according to their GRC achievement into quartile groups: struggling (Q_1), at-risk (Q_2), average (Q_3) and good readers (Q_4). The results showed that our decision to design groups of students with regard to their GRC was a sensible one.

Progress of Individual Groups of Readers in EG and CG by Summarizing

To some extent, our results support the first assumption in all three parts. First, the results showed that all students (in the EG and in the CG) made progress in summarizing, since students in both groups created summaries with better quality at the end of the intervention. Second, we can confirm the part of the assumption that students in the EG would make more extensive progress after the intervention, to a certain degree. Namely, three groups of readers – struggling, at risk and average readers in the EG made significant progress. There was also evidence of significant progress in the groups of struggling and good readers in the CG. Third, the results supported our assumption about the existence of stable proficiency differences between quartile groups of readers only in the CG of students (Pfost et al., 2014), however, we found a small decrease of the differences between the extreme groups of readers (Q_1 and Q_4) in the EG. These results indicate that we were able to make a somewhat more homogeneous group of students during the implementation of our program. This is important not only from the students' point of view, but also for the teachers, whose work

is facilitated if they teach a more homogeneous group.

In the next step, we investigated in more detail the progress of quartile groups of students in the EG, compared to the same quartile groups in the CG. Our results show that struggling readers in the EG and in the CG made significant progress during and after our intervention program, yet we found no significant differences between any of the groups, either at pre- or at the follow-up test. Consequently, with our intervention program, we were not able to improve the summarizing skill of the struggling readers in the EG to the extent that they would make significantly better summaries than this group of readers in the CG. One of the reasons might be that in both groups of struggling readers (EG and CG), students at this stage have poor reading competency. Many of them still do not have the reading technique automatized. They do not read fluently yet, which means they use a lot (too much) of mental energy from their working memory (Hintze et al., 2002). In many cases, these students have narrow vocabulary, which impedes the understanding of the reading material (Pečjak et al., 2009). Hence, our results suggest that it is too demanding for struggling readers to use summarizing strategy independently at this age (Brown et al., 1983). Most of them probably used only the copy-delete strategy when creating a summary, which means they chose a few sentences from the text and literally copied them, more or less finding main ideas only by accident.

We found a similar pattern in the groups of at-risk readers (EG and CG). There were no significant differences between them, either before the intervention or at the follow-up test. However, at-risk readers in the EG made significant progress compared to equally competent readers in the CG, but the progress was not extensive enough to make a significant difference at the end of the program. In the group of average readers, the results of students in the EG indi-

cated lower achievement in summarizing before the intervention compared to these readers in the CG, but the difference was not significant. With our intervention, average readers in the EG were able to make a shift and with significant progress created better summaries than the average readers in the CG did at the follow-up test. The achievements of at-risk and average readers imply that students, who have the reading technique at least partly automatized, were able to create better summaries after the training of summarizing skills, which was confirmed in some previous studies as well (McKenna & Stahl, 2003; Nunes et al., 2012; Pečjak et al., 2011).

The good readers in the EG and in the CG did not differ significantly in their summarizing achievement before the intervention, but according to students' results at the follow-up test, we can conclude with an 8.1% of risk rate that good readers in the EG created better summaries compared to their peers in CG, although the latter made more extensive progress. This indicates that good readers with good GRC can be trained in summarizing merely by working with texts during regular lessons, which can be explained by the fact that their entire attention is focused on creating a summary. They do not have to consider issues regarding decoding or poor understanding because of modest vocabulary etc., as in the case of the other groups of readers (Aarnoutse et al., 2001).

At the end of the program, we examined if students in all groups of readers in the EG and in the CG were able to make a transfer of their trained summarizing skill from a short to a longer text. The results showed that all groups of readers in the EG, compared to the groups in the CG, were able to make a better transfer, with the exception of the struggling readers. At-risk, average and good readers created significantly better summaries of the longer text than their peers in the CG. This indicates that these groups of students mastered the skill of summarizing to

such a degree that they were able to use it in a different learning context.

The Progress of Groups of Readers in the EG in Summarizing Elements

Finally, our goal was to examine and more thoroughly analyze the use of individual elements of summarizing by the EG before and after the intervention. Therefore, we evaluated the quality of the summaries according to the number of main ideas and the coherence of the text. Finding main ideas and connecting them into a coherent whole are two key activities, which enable students to create a good summary (Westby et al., 2010). We found significant differences between groups of students in their ability to find the main ideas, which decreased slightly in the follow-up test. This suggests that our program had a compensatory effect to some extent (Pfost et al., 2014). It seems that teachers were able to teach struggling and at-risk readers to search for (and find) most important information in the text relatively quickly by applying the program, which lasted for six months. This is important considering the fact that different students have diverse dynamics in reading skills development, and as discussed above, reading skills are a prerequisite for summarizing. Namely, struggling and at risk readers are considered to be late starters in summarizing (i.e., students with developmental delay) (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996).

We can conclude that our intervention program had the characteristics of a developmental-lag model of reading development (Parrila, Aunola, Leskinen, Nurmi, & Kirby, 2005) in finding main ideas.

With regard to the coherence of the summary, our results indicate the existence of significant differences with a large effect size between groups of students even before the intervention. The differences increased slightly

at the follow-up measurement, but the pattern of stable proficiency development could be established (Pfost et al., 2014). All groups of readers made a significant progress, which indicates that by systematically training summarizing skills we were able to develop the process of integration already in younger students. This process demands that students integrate the selected important ideas into a meaningful whole. The integration process is more common for older students in upper secondary and high school (Franzke, Kintsch, Carmicase, Johnson, & Dooley, 2005). Namely, coherence of the summary correlates with higher cognitive (executive) functions – especially with cognitive flexibility (Adams, 1990; Cartwright, 2002). For a coherent summary, the reader must switch between different language levels (syntactic, phonological and semantic) and combine them thoughtfully into a comprehensive unit. Hence, we find the significant progress of poorer readers in the EG especially important.

To sum up, our intervention program stimulated an improvement in the general ability of summarizing in all groups of readers. They included more important ideas and formed summaries that were more coherent after the intervention. However, the differences between high- and low-performing students remained. Therefore, our second assumption that better readers would make more progress than the disadvantaged ones could not be confirmed. Coherence of the summary, where we saw a trend of a more intensive progress in the group of average and good readers, was an exception.

Finally, we should emphasize that the results of our study are generalizable across different languages because the effects of training in summarizing depend on specific language and reading competency of students. Namely, how well they master the structure of their language.

Study Limitations and Pedagogical Implications

The effect of our intervention program has to be evaluated with regard to which reading skill we were developing. Reading skills differ in the degree of constraints that determine the speed of developing an individual skill. Students learn highly constrained reading skills (e.g., letter knowledge) in a shorter period, compared to less constrained skills (e.g., reading fluency), and more quickly than comprehension (comprising the summarizing skill), which is the least constrained skill (Paris, 2005). Therefore, even small training effects – especially in younger students – are very important.

In further summarizing interventions, more students should be included in each of the quartile groups and the individualization of the process of acquiring this skill for different groups of readers should be more emphasized. Especially among the struggling readers, which progressed the least, additional didactic support would be beneficial – e.g., mnemonic cards WIN (Write a topic sentence, Identify important information, and Number the FRI – facts, reasons, and ideas from the author; Saddler, Asaro-Saddler, Moeyaert, & Ellis-Robinson, 2017).

Despite the fact that we used different forms of strategy instruction in our program (approximately 50% of the training time the teachers worked frontally with the students and 50% of the time students worked in pairs or small groups), the amount of collaboration between the students should be increased because such approach enhances the students' achievement (Guthrie et al., 2004; Kyndt et al., 2013). We suggest more teacher-guided work with smaller homogeneous groups of students, which showed to be an effective way of instruction (Spörer & Brunstein, 2009).

References

- Aarnoutse, C., Van Leeuwe, J., Voeten, M., & Oud, H. (2001). Development of decoding, reading comprehension, vocabulary and spelling during the elementary school years. *Reading and Writing, 14*(1-2), 61–89.
- Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Afflerbach, P. (2016). An overview of individual differences in reading: Reader, texts, and context. In P. Afflerbach (Ed.), *Handbook of individual differences in reading* (pp. 1–12). New York, NY: Routledge.
- Baumert, J., Nagy, G., & Lehmann, R. (2012). Cumulative advantages and the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools?. *Child Development, 83*(4), 1347–1367.
- Brown, A. L., & Day, J. D. (1983). *Macrorules for summarizing texts: The development of expertise* (Technical report No. 270). Champaign, Illinois: University of Illinois at Urbana-Champaign.
- Brown, A. L., Day, J. D., & Jones, R. S. (1983). The development of plans for summarizing texts. *Child Development, 54*(4), 968–979.
- Callan, G. L., Marchant, G. J., Finch, W. H., & Flegge, L. (2017). Student and school SES, gender, strategy use, and achievement. *Psychology in the Schools, 54*(9), 1106–1122.
- Cartwright, K. B. (2002). Cognitive development and reading: The relation of reading-specific multiple classification skill to reading comprehension in elementary school children. *Journal of Educational Psychology, 94*(1), 56–63.
- Common Core State Standards (2014, February 13th). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects*. Retrieved from: http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf.
- Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99*(2), 311–325.
- Curriculum for the Slovene language (2011, February 13th). Retrieved from http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/podrocje/os/prenovljeni_UN/UN_slovenscina_OS.pdf.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3–17.
- Friend, R. (2001). Effects of strategy instruction on summary writing of college students. *Contemporary Educational Psychology, 26*(1), 3–24.
- Franzke, M., Kintsch, E., Carmicase, D., Johnson, N., & Dooley, S. (2005). Summary street: Computer support for comprehension and writing. *Journal of Educational Computing Research, 33*(1), 53–80.
- Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., ... & Tonks, S. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology, 96*(3), 403–423.
- Hattie, J. A., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning, 1*(1), 16013.
- Hershel, M. (1963). Test of Reading Level 3 – Elementary Form. M. Herschel (Ed.), *The 20th Yearbook of the National Council on Measurement in Education*. Ann Arbor, MI: SGSR.
- Hintze, J. M., Matthews, W. J., Williams, S. A. S., & Tobin, K. G. (2002). Oral reading fluency and prediction of reading comprehension in African American and Caucasian elementary school children. *School Psychology Review, 31*(4), 540–553.
- Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings?. *Educational Research Review, 10*, 133–149.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363–394.
- McKenna, M. C., & Stahl, S. A. (2003). *Assessment for Reading Instruction*. New York: Guilford.
- McKown, B. A., & Barnett, B. S. (2007). *Improving reading comprehension through higher-order thinking skills*. Research project. Chicago, IL: Saint Xavier University.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. Boston College: TIMSS & PIRLS International Study Center.
- Nation, I. S. P. (2004). Vocabulary learning and intensive reading. *EA Journal, 21*(2), 20–29.
- Nouwens, S., Groen, M. A., & Verhoeven, L. (2015). How storage and executive functions contribute to children's reading comprehension. *Learning and Individual Differences, 47*, 96–102.

- Nunes, T., Bryant, P., & Barros, R. (2012). The development of word recognition and its significance for comprehension and fluency. *Journal of Educational Psychology, 100*(4), 959–973.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117–175.
- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184–202.
- Parrila, R., Aunola, K., Leskinen, E., Nurmi, J. E., & Kirby, J. (2005). Development of individual differences in reading: Results from longitudinal studies in English and Finnish. *Journal of Educational Psychology, 97*(3), 299–319.
- Pečjak, S., & Pirc, T. (2018). Developing summarizing skills in 4th grade students: Intervention programme effects. *International Electronic Journal of Elementary Education, 10*(5), 571–581.
- Pečjak, S., Podlesek, A., & Pirc, T. (2009). Model of reading comprehension for 5th grade students. *Studia Psychologica, 51*(1), 53–68.
- Pečjak, S., Potočnik, N., & Podlesek, A. (2011). *Bralni test [Reading test]*. Ljubljana: Center za psihodiagnostična sredstva.
- Pfost, M., Hattie, J., Doerfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Mathew effects in reading. *Review of Educational Research, 84*(2), 203–244.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554–566.
- Reading Quest Organization (2017, January 12th). *Strategies for reading comprehension: Summarizing*. Retrieved from <http://www.readingquest.org/summarize.html>.
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measurements of effect size in educational research. *Educational Research Review, 6*(2), 135–147.
- Riddle Buly, M., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading tests. *Educational Evaluation and Policy Analysis, 24*(3), 219–239.
- Saarnio, D. A., Oka, E. R., & Paris, S. G. (1990). Developmental predictors of children's reading comprehension. In T. H. Carr, & B. A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 255–282). San Diego, CA, US: Academic Press.
- Saddler, B., Asaro-Saddler, K., Moeyaert, M., & Ellis-Robinson, T. (2017). Effects of a summarizing strategy on written summaries of children with emotional and behavioral disorders. *Remedial and Special Education, 38*(2), 87–97.
- Schwanenflugel, P. J., & Flanagan Knapp, N. (2016). *The psychology of reading*. New York, London: The Guilford Press.
- Spörer, N., & Brunstein, J. C. (2009). Fostering the reading comprehension of secondary school students through peer-assisted learning: Effects on strategy knowledge, strategy use, and task performance. *Contemporary Educational Psychology, 34*(4), 289–297.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, NY: Guilford Press.
- Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*(1), 77–96.
- Toličič, I., & Zorman, L. (1977). *Okolje in uspešnost učencev [Environment and students' achievement]*. Ljubljana: Državna založba Slovenije.
- Westby, C., Culatta, B., Lawrence, B., & Hall-Kenyon, K. (2010). Summarizing expository texts. *Topics in Language Disorders, 30*(4), 275–287.
- Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 53–79.