Insights into Learning Disabilities 14(2), 167-188, 2017

Improving the Length and Quality of Texts Written by Fourth Graders With Learning Disabilities Through a Peer-Tutoring Graphic Organizing Strategy

Matthias Grünke
University of Cologne, Germany

Jürgen Wilbert
University of Potsdam, Germany

Ioanna K. Tsiriotakis
University of Crete, Greece

Ainhize Lopez Agirregoikoa[1]
University of the Basque Country, Spain

A large number of students struggle with composition writing. This is alarming since problems in this area reduce the chances of succeeding in school and life in general. Children and youth with learning disabilities (LD) are especially at risk of not acquiring the skills necessary to produce texts of an acceptable length and quality. Specifically, they have difficulty planning for writing. Of the broad arsenal of instructional strategies available, graphic organizers seem to be a particularly promising tool for helping students with LD prepare to compose texts of an acceptable quality. The present study evaluated the effectiveness of a procedure whereby the use of a graphic organizer (a story map) is taught to three fourth graders with LD by three non-labeled classmates to increase the length and the quality of their text products. After only four to seven lessons, the struggling students showed stark increases in their performance. All measures of the benefits of the intervention (visual analysis, effect size measures, randomization test, and piecewise regressions) indicated that the treatment met its aim. Implications and recommendations for practice are provided.

Keywords: **Learning Disabilities, Writing Interventions, Graphic Organizer, Single-Case Study, Planning Skills.**

Introduction

Significance of Writing Skills

Writing is the principal means through which we transmit information. According to the path-breaking theory of Hayes and Flower (1980; 1986), writing is a goal-directed act of cognition that involves three core processes: planning

1 Beneficiary of the Predoctoral Formation Programme directed to Non-Doctorated Research Personnel of the Education, Linguistic Policy and Culture Department of the Basque Government.

Insights into Learning Disabilities is published by Learning Disabilities Worldwide (LDW). For further information about learning disabilities, LDW's many other publications and membership, please visit our website: www.ldworldwide.org.

(purpose, setting, generation, and organization of ideas), translating (expressing ideas in written form), and revising (correcting and rewriting). By composing meaningful text, we explore, organize, and improve our ideas (Lane et al., 2008). As such, writing competence is widely viewed as a prerequisite for school and future success, as well as an essential condition for social inclusion as it is an effective way to maintain personal links (Graham, Gillespie & McKeown, 2013; Graham & Perin, 2007).

Prerequisites for Becoming a Proficient Writer

Becoming skillful at producing text is difficult. Thus, developing writers face cognitive overload during the composition process due to having to perform multiple activities concurrently. That is, several mental resources are involved in the simultaneous process of activating prior knowledge, generating content, planning, formulating, and revising, as well as taking the intended audience and communicative purpose into account (Koster, Bouwer, & van den Bergh, 2017). Writing places especially great demands on basic cognitive processes (Archibald & Gathercole, 2006; Arfé, Dockrell, & De Bernardi, 2016; Berninger, García, & Abbott, 2009; Broc et al., 2013). Further, lower-level abilities need to reach an acceptable degree of proficiency before succeeding higher-order competencies (planning, translating, and revising) can be developed and executed. Basic skills that need to be mastered in order to be able to produce texts of relatively high inherent quality, ample lexical richness, and reasonable length include adequate orthographic knowledge (e.g., Bayat & Küçükayar, 2016; Wolf, Abbott, & Berninger, 2017) and sufficiently fluent handwriting or typing capabilities (Graham, 2010). Having to constantly think about how a word is spelled diminishes the amount of energy left to attend to the composition of the text itself. In a similar manner, difficulties with the physical process of putting thoughts on paper or into a keyboard make text production even more challenging and strenuous (Grünke, Büyüknarci, Wilbert, & Breuer, 2015).

Students with Learning Disabilities (LD) as a High-Risk Group for Failing to Meet Basic Writing Standards

Most students acquire adequate composition abilities by the end of their formal school years. However, a considerable number does not. According to the writing assessments by the National Center for Educational Statistics (2012), only about 25% of children and youth perform at the proficient level Students who are especially at risk of falling behind in writing acquisition are those with learning disabilities (LD). In a broad sense, they are characterized by failing to "… develop the knowledge, skill, will, and self-regulation necessary to succeed in key subject areas" (Grünke & Morrison Cavendish, 2016, p. 1). As a result, children and youth with LD are in grave danger of never acquiring sufficient writing competence, a factor that will impede their individual growth and general welfare (Grünke & Leonard-Zabel, 2015).

Students with LD adhere to low-level goals such as spelling, syntax errors, or even just completing a sentence. Further, they often compose texts with fewer words than their typically developing counterparts (Broc et al., 2013; Mackie & Dockrell, 2004). Moreover, they not only show less knowledge about the processes of text production compared to typically achieving writers (Scardamalia & Bereiter, 1986), they also engage in planning infrequently and ineffectively (Graham & Harris, 2007). Indeed, MacArthur and Graham (1987) found that children with LD tend to dedicate hardly any time to the act of organizing and outlining their compositions.

Importance of Providing Appropriate Writing Interventions for Students With LD

Given the lifelong importance of acquiring adequate writing skills and the challenges students with LD have in grasping fundamental text-production skills through conventional instructional writing programs, it is critical to meet their unique needs.

These learners usually require much more time and additional practice to progress and develop as writers, compared to their typical peers (Viel-Ruma, Houchins, & Fredrick, 2007). Further, even if they were able to follow a traditional writing-instruction program so that they could gain knowledge about various genres and about how to acquire text-production skills, they are rarely provided with this opportunity. Specifically, in their review of studies on writing production, Dockrell, Connelly, Walter, and Critten (2015) found that educators mainly focused on teaching basic reading and spelling abilities, rather than more complex processes, such as planning or revising. Their review also showed that many teachers found it challenging to instruct learners on how to write – a fact that may lead to detrimental effects on students' learning progress, that is, failure to acquire the necessary skills to compose texts of an adequate quality (Troia, 2010).

Efficacy of Writing Interventions for Students With LD

Luckily, a number of well-proven tools are available for teaching composition skills, including a broad database on how to support students with LD. For example, Berninger et al. (2002) pointed out that writing can be co-constructed via dyadic discussion. In particular, strategies that involve dyadic relationships (scaffolding the writing processes, modeling, or collaborative writing) have been found to improve the text quality of struggling learners (Berninger et al., 2002; Yarrow & Topping, 2001).

Peer-tutoring is the most common way of implementing dyadic relationships in educational settings. Defined as the pairing of two students (a dyad), one of whom is competent in a skill and another who is less competent, so as to enhance and extend academic instruction (Mercer, Mercer, & Pullen,

2011), peer-tutoring generally yields high benefits for struggling learners (Bowman-Perrott et al., 2013; Okilwa & Shelby, 2010).

In a comprehensive meta-analysis of writing interventions for students with LD, Gillespie and Graham (2014) found that interventions applied to foster text-composition skills reached an overall average weighted effect size (ES) of $d_{Cohen} = 0.74$. Having a skilled classmate (or another competent learner) model the use of a plan containing a series of activities designed to accomplish a writing task and then having them scaffold and support the strategy proved to be especially helpful, yielding an ES of $d_{Cohen} = 1.09$.

As mentioned, it is exceptionally difficult for students with LD to plan a text. Thus, even though they do not seem to face the same challenges with regard to performing the other two central processes (translating and revising) of Hayes and Flower's (1986) model, their ability to generate and organize ideas prior to engaging in text production is usually far from well established (Rodríguez, Grünke, Gonzalez-Castro, & Álvarez-García, 2015). Therefore, once children and youth with LD have developed adequate spelling, handwriting, and typing skills, acquisition of text-planning competencies should be targeted (Grünke & Leonard-Zabel, 2015). Since, most likely, students are initially exposed to literacy by engaging in authoring narratives, subsequently progressing to the composition of other genres, it seems logical to commence instruction by teaching them how to plan the composition of this text type.

As indicated above, interventions that include dyadic discussion (involving scaffolding, tutor modeling, and collaborative writing) have been found to be especially promising for students with LD. A meta-analysis of the effects of various prewriting activities aimed at improving the text-planning skills of students with LD revealed that most of the approaches involved the use of a graphic organizer such as a concept or a story map (Gillespie & Graham, 2014). The studies were very heterogeneous and focused on various age groups, text genres, and settings. However, only one experiment (Kurtz, 1987) examined whether the use of visual displays by children in grades four to six improved their ability to plan a narrative.

Using Peer-Tutoring to Teach Text-Planning Skills to Students With LD Through Graphic Organizers

The effectiveness of graphic organizers in improving the text-planning skills of children and youth with LD is an area gaining increasing research interest (Hennes, Büyüknarci, Rietz, & Grünke, 2015; Li, 2007; Unzueta & Barbetta, 2012; Zipprich, 1995). However, to the authors' knowledge, only two studies to date have included some form of peer-tutoring: In an experiment by Grünke, Janning, and Sperling (2016), third graders showed increased productivity in narratives written after they had received systematic support from their high-performing classmates. However, the quality of the texts was not assessed.

Using a similar kind of intervention, Grünke, Asaro-Saddler, Moeyaert, and Saddler (in preparation) found that secondary-level students also demonstrated a significant increase in the number of words written. In addition, their ability to compose narratives of reasonable quality also improved.

Purpose of the Present Study

The purpose of the present experiment was to gain further insight into the effects of graphic organizer use in teaching story-planning skills to elementary-level students with LD through peer-tutoring. In the two previous studies by Grünke and colleagues (a) the quality of the writing products was not considered (Grünke et al., 2016) and (b) the focus was on older students (Grünke et al., in preparation).

To fill these gaps, the present experiment was designed to determine whether dyadic intervention involving visual displays does not only increase the length, but also the quality of narratives produced by children with LD at the end of their elementary-school years. This aim seems meaningful, as student at this stage are expected to acquire the skill of writing acceptable stories that would, in turn, enable them to produce other genres (e.g., informative or argumentative texts) in secondary school.

Method

Participants and Setting

A convenience sample from a fourth-grade classroom in an inclusive elementary school located in a mid-sized town (58,000 citizens) in Northrhein-Westfalia, Germany, was used. The school had been collaborating with the first author on various projects for a number of years. Four hundred and ten students were enrolled in the school through grades 1 to 4. According to the principal, around 20% of them were of immigrant background.

The subtest "Spelling" from the Standardized Reading Fluency Test (German: SLRT II [Lese- und Rechtschreibtest]) by Moll and Landerl (2014) was administered to the whole class. In addition, the students were requested to write a story about a drawing that showed a conversation between a woman and a man. They were told to write a narrative that was as intriguing and compelling as possible. No time limits were set. From among the students who scored above the 50th percentile on the spelling subtest, the three students who wrote the shortest and the three students who wrote the longest texts were selected. The first three participants functioned as tutees, the last three participants served as tutors. None of the subjects had an immigrant background.

All of our tutees had been diagnosed with an LD by a multiprofessional team. They performed below average in all core school subjects and demonstrated rather low learning aptitudes. The first tutee (Hendrik) was a 9-year-old boy. According to his classroom teacher, he experienced difficulties in math and

in acquiring new information. The second tutee (Anja) was an 11-year-old girl. Her classroom teacher reported that she faced severe learning difficulties and exhibited considerable concentration problems. The third tutee, 9-year-old Ilja, was male. He had demonstrated speech, language, and communication difficulties (SLCD) as a preschooler, but his classroom teacher pointed out that he had made noticeable progress in the skills relevant to SLCD over the last couple of years. She noted that Ilja demonstrated oppositional defiant behavior at times. (All tutee names were changed for anonymity.)

The group of tutors consisted of two 10-year-old girls and one 9-year-old boy. According to their mid-term report cards, these students belonged to the top quarter of their class in language arts. Their classroom teacher not only described all tutors as responsible and socially skilled, but also thought all six children were able to work well together.

Dependent Variables and Measurement

Writing prompts. We used 48 paper strips with different story headlines as writing prompts, adapted from Hosp, Hosp, and Howell (2016). At every measuring point, each tutee drew two paper strips from a box that contained all the prompts that she or he had not yet seen.

Number of total words written (TWW). TWW was used to capture the quantity of the written products. This widely applied index is defined as the number of recognizable words written regardless of spelling or context (excluding digits). It has been recommended for primary use in elementary schools and for optional use at the secondary level (Malecki, 2008). According to Amato and Watkins (2001), TWW is the strongest measure correlated with a variety of writing criteria. However, it is focused on fluency and may not always be sensitive to other aspects of text production (Aitken & Martinussen, 2013).

The Scale for Scoring the Inclusion and Quality of the Parts of a Story (SIQS). We used the SIQS by Harris and Graham (1996) to assess how well each writing product was composed. It includes eight criteria (main character, locale, time, starter event, goal, action, ending, and reaction). A narrative is awarded points for incorporating the different aspects into the text. The SIQS provides explanations and examples for the various scoring options. Total number of points ranges between 0 and 20.

Accuracy and interrater reliability. To ensure accuracy, TWW was counted twice by a female student assistant, who was blind to the study. If the first two counts were different, she counted a third time. Fourth counts were never necessary. The female student assistant and another (male) student assistant were trained by the primary researcher to use the SIQS in appraising the quality of the texts. They assessed the narratives independently and subsequently compared their results. If there was a disagreement, they went back to the particular text and discussed any discrepancies until consensus was reached.

However, the overall scores turned out to be very similar (interrater reliability of over 90%, according to Richard, Taylor, & Ramasamy, 2013) even before they were harmonized.

Experimental Design and Procedure

We applied a multiple baseline across participants design with repeated measures during the maintenance phase (ABC; see Kazdin, 2011). Baseline (A) and intervention (B) lasted 13 school days. Three additional probes were designated for collecting followup data (C) on Monday, Wednesday, and Friday of the week following Phase B. Performance was measured on each of those days at the end of the third period. This period was scheduled for the treatment to take place during the intervention phase. Phase A served as control and was, therefore, compared with the B and the C Phases. We expected significant improvements in performance from Phase A to Phase B in all tutees and a consolidation of the performance level (comparable to the one at the end of Phase B) during the followup condition.

In accordance with single-case reporting guidelines in behavioral interventions (see Tate et al., 2016), we incorporated a randomization procedure, which meant that the beginning of the treatment was determined by chance within pre-defined ranges: Baseline and intervention were conducted for at least three days each. Thus, the treatment could have started after the 3rd, 4th, 5th, 6th, 7th, 8th, 9th, or 10th measurement point. A random drawing of one option for each student out of these eight alternatives (using paper slips and a small basket) yielded a starting point for Student A after the sixth probe, for Student B after the ninth probe, and for Student C after the eighth probe.

During baseline, a research assistant handed each tutee a writing prompt in the form of a paper strip with a story headline (see above). The three participants were asked to compose a narrative that was as engrossing and enjoyable as possible. They were provided with pen and paper. The research assistant told them to take as much time as they wanted to write their text. None of the students ever needed more than 10 minutes to finish the assignment.

During the intervention phase, the performance continued to be measured as during baseline. However, before administering the probes, the three tutor-tutee teams worked individually on improving the writing skills of the tutees for 30 minutes in a quiet corner of the classroom. Prior to the treatment, the tutors received three 1-hour training sessions by a female graduate university student of special education on how to best utilize the story mapping strategy with their team partners.

Every peer-tutorial session started with a presentation of a performance graph that illustrated the trajectory of the Tutee's productivity scores (TWW) up to the given day. Subsequently, the tutors gave their classmates feedback based on attribution theory (Heider, 1958; Kelley, 1967), as follows: If a tutee showed

a decrease in TWW in the measurement the day before, the tutor provided a variable explanation for the outcome ("Yesterday, you might not have tried as hard as usual," or "Maybe you had a rough day – but let's leave this behind and give it another go"). In case a tutee remained on the same level or outperformed her or his previous score, the tutor gave a feedback in which she or he emphasized internal reasons for the previous day's results ("You really tried hard and it paid off," "You applied the strategy that we are working with and thus you produced great achievements").

The tutor then provided explicit instruction on how to apply the story-mapping strategy using modeling of the desired skills (Hollingsworth & Ybarra, 2009; Reid, Lienemann, & Hagaman, 2013). The procedure consisted of the following steps: (a) Contemplate on the story heading, (b) think about what could happen in the story, (c) review the fields of the story map, and (d) take notes on your ideas about the story using the appropriate fields. An 11x17 inch poster on the wall served as a reminder of these four actions.

During the first training session, the tutors introduced the steps of the strategy while referring to the poster. They promised their classmates that applying story mapping would result in better texts and, thus, a sense of achievement. Subsequently, the tutors demonstrated the procedure in response to a writing prompt. Each prompt was randomly drawn from the pool of story headings that was used to measure the treatment effects. (Under no circumstances was a tutee presented with the same prompt more than once.)

The first two training sessions were dedicated to the tutors demonstrating the use of the story-mapping strategy and to the tutees executing it themselves while the tutors provided scaffolded feedback. During each of the first two lessons, the tutors filled out one or two story maps and the tutees filled out two or three. In the following sessions, support from the tutors was gradually faded, providing the tutees with opportunities for independent practice. From the fourth session onward, tutees were encouraged to draw their own story maps instead of using a ready-made template.

The followup phase (C) resembled the baseline phase (A). Tutees were asked to produce an exciting story without receiving any instructions or feedback. Again, they were provided with pen and paper. But whereas they did not seem to know how to make effective use of the note paper during the first days of the experiment, they started to create their own story maps by the end of the intervention and continued to do so during followup.

Fidelity of Implementation

To enhance treatment fidelity, we created a simple checklist that covered the basic features of the peer-tutoring intervention as described above. The graduate student who administered the probes sat in during each training session to make sure the intervention was implemented as intended. In case

the students got sidetracked, she stepped in and encouraged them to attend to their assignments.

Results

A visual inspection of the data for TWW (see Figure 1) shows a clear increase throughout all measurements for all three cases, with the development for Ilja the least pronounced. All three baselines can be considered as stable. An increase only occurred at the start of Phase B. For the SIQS scores (see Figure 2), the trajectory is similar, although less pronounced. Here it looks like an overall increase of the scores at the beginning of Phase B that remains stable at Phase C. Compared to the plots of the TWW, this boost is more immediate than continuous.

Figure 1. **TWW scores for all three cases.**

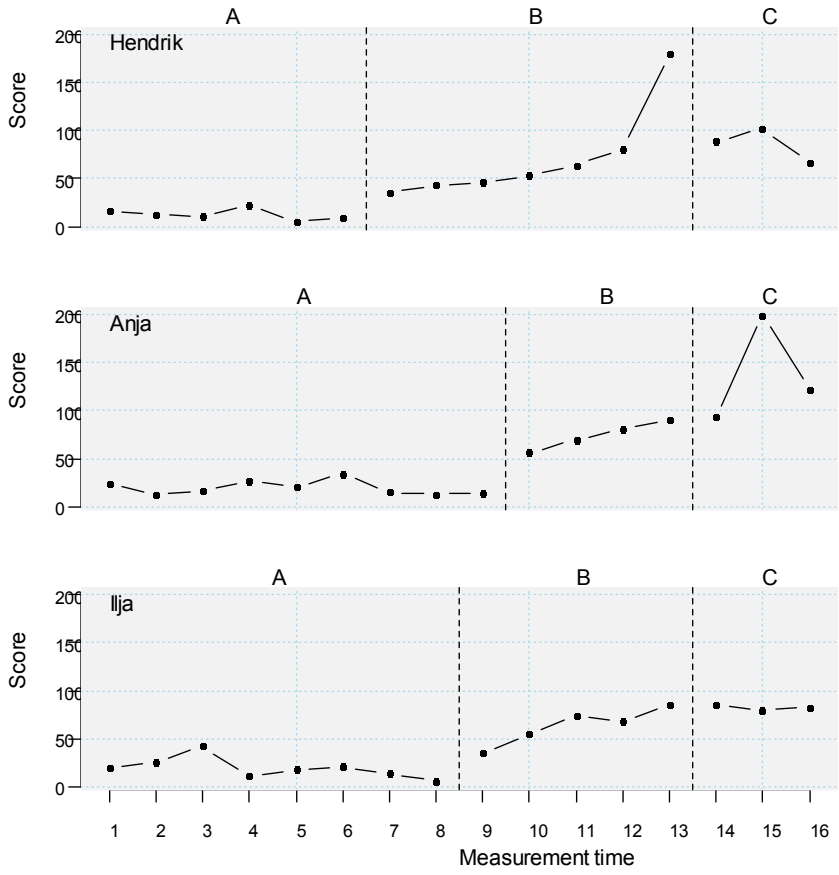

Figure 2. **SIQS scores for all three cases.**

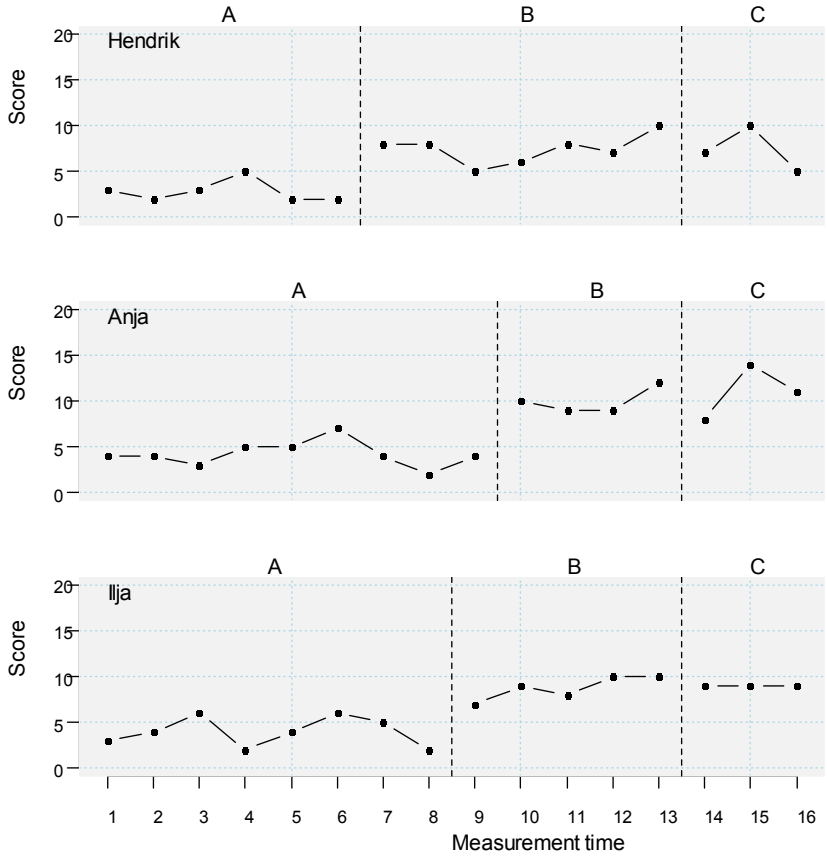

Table 1 shows the descriptive statistics of total words written (TWW) for each of the three single cases. The number of measurements (*N*), mean (*M*), median (*Md*), minimum value (*min*), maximum value (*max*), and standard deviation (*SD*) are presented separately for Phases A, B, and C. Additionally, a linear regression was calculated to examine the development across time for each case. The parameter *B* indicates changes in TWW per measurement.

As illustrated, all three cases show a stark average increase (mean and median) from Phase A to Phase B, which stabilizes in a further increase to Phase C. Furthermore, all three cases show a slight decrease of TWW within Phase A (see *B* parameter) and a clear increase in Phase B. For Phase C, the data are ambiguous.

Table 1. ***Descriptive Statistics for TWW of Each Phase***

	N			*M*			*Md*			*min*			*max*			*SD*			*B*		
Name	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
Hendrik	6	7	3	12.8	71.7	85.7	11.5	53	89	5	35	66	22	180	102	5.9	50.1	18.2	-1.3	18.9	-11.5
Anja	9	4	3	19.7	74.3	138.0	17	75	122	12	56	93	34	91	199	7.4	14.9	54.8	-0.6	11.5	14.5
Ilja	8	5	3	19.8	63.4	82.7	19	68	83	6	35	80	42	85	85	10.9	19.2	2.5	-2.5	11.3	-1.0
Total[1]	23	16	9	17.4	69.8	102.1	17	68	89	5	35	66	42	180	199	8.4	32.1	33.4	-1.5	13.9	0.7

[1]The total is calculated as the mean of the three participants except for *N*, which is the sum of the three participants.

Table 2. ***Descriptive Statistics for SIQS of Each Phase***

	N			*M*			*Md*			*min*			*max*			*SD*			*B*		
Name	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
Hendrik	6	7	3	2.8	7.4	7.3	2.5	8	7	2	5	5	5	10	10	1.2	1.6	2.5	-0.1	0.3	-1.0
Anja	9	4	3	4.2	10.0	11.0	4	9.5	11	2	9	8	7	12	14	1.4	1.4	3.0	0.0	0.6	1.5
Ilja	8	5	3	4.0	8.8	9.0	4	9	9	2	7	9	6	10	9	1.6	1.3	0.0	0.0	0.7	0.0
Total[1]	23	16	9	3.7	8.7	9.1	4	9	9	2	5	5	7	12	14	1.4	1.5	2.3	0.0	0.5	0.2

[1]The total is calculated as the mean of the three participants except for *N*, which is the sum of the three participants.

The pattern is similar for the SIQS measures (see Table 2). For each case, the mean and median values increase from Phase A to Phase B and remain stable in Phase C. While all cases show no increase in SIQS scores throughout time within Phase A, the scores are increasing in Phase B for all cases, with an average of 0.5 points per measurement. Again, the results are ambiguous for Phase C.

The next analysis focused on overlapping indices, comparing Phase A to Phase B. Seven of the most common effect sizes are reported here: percent of nonoverlapping data (PND), percent of data points exceeding the median (PEM), percent of data points exceeding the median trend (PET), nonoverlap of all pairs (NAP), nonoverlap of all pairs rescaled (NAP rescaled), percent of all nonoverlapping data (PAND), and TA-U (see Alresheed, Hott, & Bano 2013; Parker, Vannest, & Davis, 2011; Parker, Vannest, Davis, & Sauber, 2011).

Table 3 contains the data for each of the three cases and for both measures (TWW and SIQS). All indices point towards a strong effect of the intervention (i.e., an increase of the measured values) for both dependent variables (a value of 80 or higher for the first six measures is considered to indicate a strong effect; no similar clear conventions exist for Tau-U).

Table 3. Overlap Indices for TWW and SIQS

	PND	*PEM*	*PET*	*NAP*	*NAP rescaled*	*PAND*	*TAU_U*
				TWW			
Hendrik	100	100	100	100	100	100	.86
Anja	100	100	100	100	100	100	.58
Ilja	80	100	100	98	95	85	.74
				SIQS			
Hendrik	86	100	100	99	98	92	.77
Anja	100	100	100	100	100	100	.51
Ilja	100	100	100	100	100	100	.57

We also conducted a randomization test for single-case designs (Dugard, File, & Todman, 2012). Randomization tests allow for analyzing the conjoint effect of all cases, they are robust against serial dependent data, and they provide probability values for generalizing the results. The randomization test was set up in accordance with the design, allowing for at least three measurements per phase.

As depicted in Table 4, the mean difference in TWW between Phase A and Phase B did not reach statistical significance ($\Delta M_{AB} = 52.4$, $p < .11$), whereas the difference in SIQS between Phase A and Phase B was significant ($\Delta M_{AB} = 5.1$, $p < .01$). Table 4 also shows the results of a comparison between Phase A and Phase C. (This should be interpreted cautiously as there was no actual random variation at the start of the C Phase.) For both measures the mean differences were significant (TWW: $\Delta M_{AC} = 84.7$, $p < .01$; SIQS: $\Delta M_{AC} = 5.4$, $p < .01$).

Table 4. ***Randomization Test for the Mean Difference of Phases A and B or Phases A and C***

	p-level	
Variable	A vs. B[1]	A vs. C[2]
TWW	.11	.01
SIQS	.01	.01

[1]Possible permutations included the start of Phase B between measurement time 4 and 11. This resulted in 512 combinations. [2]Possible permutations were derived with the constraint that each phase has a length of at least three measurements. This resulted in 168 combinations.

The final analyses are based on piecewise regressions (Huitema & McKean, 2000). They allow for controlling developmental trends in the data (trend effects) and differentiate between continuous (slope effect) and immediate (level effects) intervention effects. Table 5 shows the results for TWW. For all three cases, there was either a significant slope, a level effect, or both. The slope effects were especially strong, with R^2 ranging between 7% and 19%.

Table 5. ***Piecewise Regression Model for TWW***

	B	*SE*	*t*	*p*	R^2
			Hendrik		
Intercept	17.33	22.39	0.77	.46	
Trend	-1.29	5.75	-0.22	.83	.00
Level	-13.48	26.76	-0.5	.63	.01
Slope	20.18	7.33	2.75	.02*	.17
			Anja		
Intercept	22.5	5.01	4.49	.01**	
Trend	-0.57	0.89	-0.64	.54	.00
Level	28.1	9.45	2.97	.02*	.04
Slope	12.07	3.21	3.76	.01**	.07
			Ilja		
Intercept	31.21	7.18	4.34	.01**	
Trend	-2.55	1.42	-1.79	.11	.03
Level	18.67	11.36	1.64	.13	.03
Slope	13.85	3.24	4.27	.01**	.19

The results for the SIQS score are less clear. Hendrik showed a significant level effect ($B = 3.81$, $p < .05$, $R^2 = .13$). Anja's level increase was even larger ($B = 4.41$, $p = .06$, $R^2 = .09$) but failed to be significant due to a larger variance. Finally, for Ilja, neither slope nor level effects were significant.

Table 6. ***Piecewise Regression Model for SIQS***

	B	*SE*	*t*	*p*	R^2
			Hendrik		
Intercept	3.13	1.41	2.22	.05*	
Trend	-0.09	0.36	-0.24	.82	.00
Level	3.81	1.69	2.26	.05*	.13
Slope	0.34	0.46	0.73	.49	.01
			Anja		
Intercept	4.39	1.07	4.08	.01**	
Trend	-0.03	0.19	-0.17	.87	.00
Level	4.41	2.03	2.18	.06	.09
Slope	0.63	0.69	0.92	.38	.02
			Ilja		
Intercept	4.00	1.16	3.45	.01**	
Trend	0.00	0.23	0.00	1.0	.00
Level	2.70	1.83	1.47	.17	.05
Slope	0.70	0.52	1.34	.21	.04

Discussion

Basic Findings

Few studies to date have focused on the effects of using graphic organizers to teach story-planning skills to elementary school students with LD through peer-tutoring strategies. The aim of the present experiment was to extend previous findings to help overcome this significant limitation.

Overall, the results of this single-case analysis indicate a functional relationship between the onset of the intervention and an increase in writing performance. During the course of between four to seven highly structured peer-tutorial sessions, all three tutees were able to compose markedly longer and qualitatively better narratives. In addition, they displayed increased post-intervention performance.

Visual analysis of the data suggests that the treatment had a positive effect on the writing competence of the students with LD participating in this

study. This appraisal was backed up by a number of overlap indices, which quantified the differences between Phases A and B. All of them reached a magnitude that speaks to the effectiveness of the intervention. Although a randomization test did not reveal a statistically significant difference between the means of TWW in Phases A and B, the discrepancies between the average SIQS scores across Phases A and B were statistically significant ($p < .01$).

Besides, piecewise regressions detected slope or level effects (or both) concerning text length in all three tutees. However, with one exception, the piecewise regressions did not verify statistically significant improvements in text quality. Only Hendrik showed a momentous difference in level between Phases A and B. But on the whole, there is no reasonable doubt that the dyadic intervention was effective not only in enhancing the length, but also the quality of the narratives written by the fourth graders who participated in this experiment.

As such, these findings support the limited body of previous research regarding the potential of using peer-tutoring as a means of teaching text planning skills to students with LD. All existing results support confidence in this approach and give rise to the hope that writing problems can be effectively treated.

Limitations

Despite the promising results, the study is not exempt from limitations. Initially, caution must be used in generalizing the results. Further research is needed to corroborate the effects and substantiate the potency of the peer-tutoring strategies examined as the tutees belonged to a particular age group. Thus, any conclusions based on these results can only be drawn in reference to this specific population. Additionally, broad implications about the benefits of the treatment cannot be made, as only a small sample of students participated. Hence, more research is necessitated about supporting struggling (LD) learners through peer-tutoring strategies to assist them in improving their writing-planning skills.

A further limitation, the results pertain to the specific text genre investigated. Hence, any conclusions can only refer to this specific text genre. Nonetheless, it would be interesting for peer-tutoring strategies to be further implemented on other text genres (e.g., informative or argumentative) and/or incorporate different samples of students across age ranges and abilities.

Even though the study design included a follow-up condition to determine maintenance effects, three measurement points are not enough to justify the conclusion that the treatment has lasting positive consequences on the writing performance of fourth graders with LD.

A final limitation pertains to the nature of capturing text length and quality. Measuring writing is challenging, because no set of best practices exists in this respect (Van Steendam, Tillema, Rijlaarsdam, & Van den Bergh, 2012). Counting the number of TWW is not the only option to quantify productivity. Other prominent options include correct word sequences (CWS), correct minus

incorrect word sequences (CIWS), or correctly spelled words (CSW) (Dockrell et al., 2015; Hosp et al., 2016; McMaster et al., 2017). Each of these alternatives has its advantages, but every approach is also subject to criticism on a number of accounts (Gansle, Van Der Heyden, Noell, Resetar, & Williams, 2006; Hampton & Lembke, 2016).

The same applies to the many strategies of capturing text quality. Applying writing rubrics is the most common way to monitor the level of excellence of writing products over time. However, the sheer number of these scoring guides is confusing. For example, almost each single-case study in the meta-analyses by Cook and Bennett (2014), Gillespie and Graham (2014), and Rogers and Graham (2008) used its own self-constructed tool for assessing performance.

By reverting to TWW and an already existing scale (SIQS), we tried to incorporate instruments that are relatively common and familiar to most researchers working in the field of writing.

Implications for Practice

Despite these limitations, the findings of this study have several valuable implications for practitioners working with children with LD, showing that teachers can embed writing-planning instructions into their everyday classroom activities without much effort. The importance of helping students with LD to acquire adequate writing skills so that they can subsequently engage in the production of other types of texts (e.g., informative or argumentative) in secondary school remains vital. The findings derived from this investigation have implications for the writing instruction of elementary students with LD, as well as for these students' progression in secondary school.

Peer-tutoring strategies could be more beneficial in inclusive classrooms settings than other strategies used to improve writing as they focus on fostering a safe and collaborative environment where the dyadic tutor-tutee relation established via feedback and assistance is basically at the very core of this procedure. That is, the dyadic tutor-tutee union established among students could prove to be a dynamic factor in enhancing students' writing motivation while making the process less tedious for them. Through the tutor-tutee union, children could view the act of composing as a positive teamwork activity rather than a demanding, rigid task. Further, engaging all students in tutor role-play supports inclusion inside a school. Additionally, peer-tutoring strategies can effectively and efficiently be implemented and carried out in general education classrooms, creating a student- rather than a teacher-centered environment.

Conclusion

The results of this study highlight that peer-tutoring is a significant tool for general education classrooms as it not only improves the text-production skills of children with LD, but also strengthens collaboration and partnership between students of different levels and abilities. The No Child Left Behind

policy calls upon teachers to meet the needs of all students. An integral part of helping learners to develop higher-level thinking and content knowledge aligned with the Common Core State Standards could be through the use of story maps and peer-tutoring strategies. This would allow students with LD to improve their text-production skills in a highly engaging, collaborative environment and be sufficiently prepared for secondary education.

As mentioned above, future research should focus on replicating the results of this study and addressing the previously listed limitations. Moreover, it would be interesting to investigate and explore the benefits of the strategy implemented here on the students acting as tutors.

References

Aitken, M., & Martinussen, R. (2013). Exploring predictors of performance on a curriculum-based measure of written expression, *Journal of Writing Research, 4*, 281-299.

Alresheed, F., Hott, B. L., & Bano, C. (2013). Single subject research: A synthesis of analytic methods. *The Journal of Special Education Apprenticeship, 2*, 1-18.

Amato, J. M., & Watkins, M. W. (2011). The predictive validity of CBM writing indices for eighth grade students. *The Journal of Special Education, 44*, 195-204.

Archibald, L. M. D., & Gathercole, S. E. (2006). Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. *Psychonomic Bullerin & Review, 14*, 919-924.

Arfé, B., Dockrell, J. E., & De Bernardi, B. (2016). The effect of language specific factors on early written composition: the role of spelling, oral language and text generation skills in a shallow orthography. *Reading and Writing, 29,* 501-527.

Bayat, S., & Küçükayar, H. (2016). Relationship between legible handwriting and level of success of third grade students in written expression. *Journal of Education and Training, 4*, 33-41.

Berninger, V., Fayol, M. L., & Alamargot, D. (2012). Learning to spell words with the pattern analyser, oracle, cross-code talker, cross-code scribe, and silent orthographer. In M. Fayol, D. Alamargot, & V. W. Berninger (Eds.), *Translation of thought to written text while composing* (pp. 71-93). New York, NY: Taylor & Francis.

Berninger, V. W., García, N., & Abbott, R. (2009). Multiple processes that matter in the writing instruction and assessment. In G. Troia (Ed.), *Instruction and assessment for struggling writers: Evidence based practices* (pp. 15-50). New York, NY: Guilford.

Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., Graham, S., & Hawkins, J. M. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal od Educational Psychology, 94*, 291-304.

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer-tutoring: A meta-analytic review of single-case research. *School Psychology Review, 42*, 39-55.

Broc, L., Bernicot, J., Olive, T., Favart, M., Reilly, J., Quémart, P., & Uzé, J. (2013). Lexical spelling in children and adolescents with specific language impairment: Variations with the writing situation. *Research in Developmental Disabilities, 34,* 3253-3266.

Cook, K. B., & Bennett, K. E. (2014). Writing interventions for high school students with disabilities: A review of single-case design studies. *Remedial and Special Education, 35*, 344-355.

Dockrell, J. E., Connelly, V., Walter, K., & Critten, S. (2015). Assessing children's writing products: The role of curriculum based measures. *British Educational Research Journal, 41*, 575-595.

Dugard, P., File, P., & Todman, J. (2012). *Single-case and small-n experimental designs: A practical guide to randomization tests*. New York, NY: Routledge.

Gansle, K. A., Van Der Heyden, A. M., Noell, G. H., Resetar, J. L., & Williams, K. L. (2006). The technical adequacy of curriculum-based and rating-based measures of written expression for elementary school students. *School Psychology Review, 35*, 435-450.

Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children, 80*, 454-473.

Graham, S. (2010). Want to improve children's writing? Don't neglect their handwriting. *Education Digest, 76*, 49-55.

Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: importance, development and instruction. *Reading and Writing, 26*, 1-15.

Graham, S., & Harris, K. R. (2007). Best practices in teaching planning. In S. Graham, C. A. MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (pp. 119-140). New York, NY: Guilford.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*, 445-476.

Grünke, M., Asaro-Saddler, K., Moeyaert, M., & Saddler, B. (in preparation). The effects of a peer-tutoring intervention on the text productivity and quality of narratives written by eighth graders with learning disabilities.

Grünke, M., Büyüknarci, Ö., Wilbert, J., & Breuer, E. (2015). To what extent do certain characteristics of a child's written story influence the way it is rated? *Insights Into Learning Disabilities, 12*, 163-177.

Grünke, M., Janning, A. M., & Sperling, M. (2016). The effects of a peer-tutoring intervention on the text production of students with learning and speech problems: A case report. *Learning Disabilities: A Contemporary Journal*, *14*, 225-235.

Grünke, M., & Leonard-Zabel, A. M. (2015). How to support struggling writers: What the research stipulates. *International Journal of Special Education*, *30*, 137-149.

Grünke, M., & Morrison Cavendish, W. (2016). Learning disabilities around the globe: Making sense of the heterogeneity of the different viewpoints. *Learning Disabilities: A Contemporary Journal, 14*, 1-8.

Hampton, D. D., & Lembke, E. S. (2016). Examining the technical adequacy of progress monitoring using early writing curriculum-based measures. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, *32*, 336-352.

Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation*. Brookline, MA: Brookline Books.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Greg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum.

Hayes, J. R., & Flower, L. S. (1986). Writing research and the writer. *American Psychologist, 41*, 1106-1113.

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.

Hennes, A.-K., Büyüknarci, Ö., Rietz, C., Grünke, M. (2015). Helping children with specific learning disabilities to improve their narrative writing competence by teaching them the use of the story maps strategy. *Insights into Learning Disabilities, 12*, 35-56.

Hollingsworth, J. R., & Ybarra, S. E. (2009). *Explicit direct instruction (EDI): The power of the well-crafted, well-taught lesson*. Thousand Oaks, CA: Corwin.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2016). *The ABCs of CBM: A practical guide to curriculum-based measurement*. New York, NY: Guilford.

Huitema, B. E., & McKean, J. W. (2000). Design specification issues in time-series intervention models. *Educational and Psychological Measurement*, *60*, 38-58.

Kazdin, A. E. (2011). *Single-case research designs*. New York, NY: Oxford University Press.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.*), Nebraska Symposium on Motivation* (Volume 15, pp. 192-238). Lincoln, NE: University of Nebraska Press.

Koster, M., Bouwer, R., & van den Bergh, H. (2017). Professional development of teachers in the implementation of a strategy-focused writing intervention program for elementary students. *Contemporary Educational Psychology*, *49*, 1-20.

Kurtz, H. (1987). *Effects of a specific strategy training on the written expression of learning disabled students grades 3-6* (Unpublished doctoral dissertation). Drake University, Des Moines, IA.

Lane, K., Harris, K. R., Graham, S., Weisenbach, J. L., Brindle, M., & Morphy, P. (2008). The effects of self-regulated strategy development on the writing performance of second-grade students with behavioral and writing difficulties. *The Journal of Special Education, 41*, 234-253.

Li, D. (2007). Story mapping and its effects on writing fluency and word diversity of students with learning disabilities. *Learning Disabilities: A Contemporary Journal, 5*, 77-93.

MacArthur, C. A., & Graham, S. (1987). Learning disabled students' composing under three methods of text production: Handwriting, word processing and dictation. *Journal of Special Education, 21,* 22-42.

Mackie, C., & Dockrell, J. E. (2004). The nature of written language deficits in children with SLI. *Journal of Speech Language and Hearing Research, 47*, 1469-1483.

Malecki, C. (2008). Best practices in written language assessment and intervention. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 477-488). Bethesda, MD: National Association of School Psychologists.

McMaster, K. L., Shin, J., Espin, C. A., Jung, P.-G., Wayman, M. M., & Deno, S. L. (2017). Monitoring elementary students' writing progress using curriculum-based measures: Grade and gender differences. *Reading and Writing, 30*, 2069-2091.

Mercer, C. D., Mercer, A. R., & Pullen, P. C. (2011). *Teaching students with learning problems*. Upper Saddle River, NJ: Prentice Hall.

Moll, K., & Landerl, K. (2014). *Standardized Reading Fluency Test*. Bern, Switzerland: Huber.

National Center for Education Statistics. (2012). *The nation's report card: Writing 2011*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Okilwa, N. S. A., & Shelby, L. (2010). The effects of peer-tutoring on academic performance of students with disabilities in grades 6 through 12: A synthesis of the literature. *Remedial and Special Education, 31*, 450-463.

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, *35*, 303-322.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*, 284-299.

Reid, R., Lienemann, T. O., & Hagaman, J. L. (2013). *Strategy instruction for students with learning disabilities*. New York, NY: Guilford.

Richard, S. B., Taylor, R. L., & Ramasamy, R. (2013). *Single-subject research: Applications in educational and clinical settings*. Boston, MA: Wadsworth.

Rodríguez, C., Grünke, M., González-Castro, P., García, T., & Álvarez-García, D. (2015). How do students with attention deficit/hyperactivity disorders and writing learning disabilities differ from their non-labeled peers in the ability to compose texts? *Learning Disabilities: A Contemporary Journal, 13*, 157-175.

Rogers, L. A., & Graham, S. (2008). A meta-analysis of single-subject design writing intervention research. *Journal of Educational Psychology, 100*, 879-906.

Scardamalia, M., & Bereiter, C. (1986). Research on written composition. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 778-803). Basingstoke, UK: Macmillan.

Tate, R. L. et al. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016 statement. *Neuropsychological Rehabilitation, 27*, 1-15.

Troia, G. A. (2010). *Instruction and assessment for struggling writers.* New York, NY: Guilford.

Unzueta, C., & Barbetta, P. M. (2012). The use of computer graphic organizers for persuasive writing by Hispanic students with specific learning disabilities. *Journal of Special Education Technology, 27*, 15-30.

Viel-Ruma, K., Houchins, D., & Fredrick, L. (2007). Error self-correction and spelling: improving the spelling accuracy of secondary students with disabilities in written expression. *Journal of Behavioral Education, 16,* 291-301.

Van Steendam, E., Tillema, M., Rijlaarsdam, G., & Van den Bergh, H. (Eds.), (2012). *Measuring writing: Recent insights into theory, methodology and practices.* Leiden, NL: Brill.

Wolf, B., Abbott, R. D., & Berninger, V. W. (2017). Effective beginning handwriting instruction: Multi-modal, consistent format for 2 years, and linked to spelling and composing. *Reading and Writing, 30*, 299-317.

Yarrow, F., & Topping, K. J. (2001). Collaborative writing: The effects of metacognitive prompting and structured peer interaction. *British Journal of Educational Psychology, 71,* 262-282.

Zipprich, M. A. (1995). Teaching web making as a guided planning tool to improve student narrative writing. *Remedial and Special Education, 16*, 3-15.

Authors' Note

For correspondence regarding this article, please contact: Matthias Grünke, Ph.D., Department of Special Education & Rehabilitation, University of Cologne, Klosterstr. 79b, Cologne, Northrhine-Westfalia, 50931, Germany, Phone: 0049-221-4705547, Email: matthias.gruenke@uni-koeln.de.