The Effects of a Peer-Tutoring Intervention on the Text Productivity and Completeness of Narratives Written by Eighth Graders With Learning Disabilities

Matthias Grünke,[1] Bruce Saddler
Kristie Asaro-Saddler, and
Mariola Moeyaert[2]

Abstract

Writing becomes increasingly more imperative across all content areas as students progress through secondary school. However, many of them struggle with the complex process of putting thoughts and ideas onto paper or into a keyboard. Adolescents with learning disabilities (LDs), in particular, are usually challenged by writing activities. One major cause for their difficulties is their lack of planning skills. This single-case study evaluated a peer-tutoring approach designed to teach adolescents with LDs to better plan narratives by using a simple strategy (story mapping). A multiple-baseline design (AB) across subjects was employed to assess the effects of the intervention. The results indicate large and significant effects of the intervention on the length and completeness of the narratives the students produced. This suggests that writing skills of adolescents with LDs can be improved by way of peer tutoring with limited resources required.

Keywords: writing difficulties, learning disabilities, peer tutoring, writing planning skills

Significance of Writing and Causes for Problems in Text Production

Written language is a critical element in the lives of millions of students enrolled in K–12 schools (National Center for Education Statistics [NCES], 2012). Proficiency in text production is an important skill for elementary-aged children to learn, but writing effectively becomes even more essential for youth in middle and high school as they prepare to make the critical transition to higher education and/or the workforce. In school, writing is often used as an instrument to assess students' understanding of content (Graham, 2008). Text production is also an important means of sharing and communicating in everyday life through social media outlets, emailing, and texting (Graham & Hall, 2016), with an estimated 171 billion e-mail messages sent daily (NCES, 2012). In the workplace, writing is essential to acquiring employment and advancing in a profession, with more workers than ever before required to create reports, PowerPoint presentations, e-mails, or other types of written documentation (National Commission on Writing, 2004).

However, despite the centrality of written expression in learning, social, and work environments, a great percentage of school-aged children struggle with this skill. In fact, according to the most recent National Assessment of Educational Progress (NCES, 2012), of the three possible skill levels (basic, proficient, and advanced), only about a quarter (27%) of grade 8 and 12 students performed at or above a "proficient" level, with only 5% of students with disabilities performing at or above that standard. The

[1]University of Cologne, Germany; [2]University at Albany, USA

Grünke, M., Saddler, B., Asaro-Saddler, K., & Moeyaert, M. (2019). The Effects of a Peer-Tutoring Intervention on the Text Productivity and Completeness of Narratives Written by Eighth Graders With Learning Disabilities. *International Journal for Research in Learning Disabilities,* 4(1), 41-58. https://doi.org/10.28987/ijrld.4.1.41

present study was conducted in Germany. Unfortunately, no comparable surveys have been undertaken in this country on the expressive writing competencies of elementary and/or secondary students.

Perhaps one reason for the difficulties that many learners experience in writing is that they do not spend sufficient time developing writing skills and engaging in writing tasks. For example, secondary school students spend little time writing or being taught how to write (Graham, Capizzi, Harris, Herbert, & Morphy, 2014), despite a large majority of teachers believing that writing is very important beyond high school (Kiuhara, Graham, & Hawken, 2009). Apparently, what time is devoted to writing is not well spent. In one study, Rietdijk, van Weijen, Janssen, van den Bergh, and Rijlaarsdam (2018) documented that communicative writing, process writing, and writing strategy instruction were insufficiently implemented in primary education classrooms. Another survey, by Applebee and Langer (2011), revealed that secondary school students invested the majority (over 80%) of their writing time in engaging in tasks such as completing fill-in-the-blank assignments, short-answer exercises, and copying information from teacher presentations, rather than composing extended writing pieces, such as stories or essays. Such low-level writing tasks often require little analysis and interpretation (Kiuhara et al., 2009) and do not encourage students to think critically, a skill required for effective writing that is highly valued by employers (Graham & Hall, 2016; National Association of Colleges and Employers, 2012; National Commission on Writing, 2004).

Students' lack of proficiency in writing may also be related to teachers' lack of knowledge regarding high-quality writing instruction. Writing is a difficult skill to learn and to teach, and quality pre- or inservice writing instruction must first be experienced by teachers before they can deliver it to students. However, educators may not be receiving the training they need (Grünke & Leonard-Zabel, 2015). For example, Graham et al. (2014), in a survey, found that nearly two-thirds (64%) of middle school teachers reported having received minimal or no preservice training on how to teach writing. A similar lack of training may also have contributed to secondary school teachers reporting only infrequently using evidence-based writing practices and adaptations for their struggling writers (Graham et al., 2014; Kiuhara et al., 2009), an outcome that is especially problematic for students with learning disabilities (LDs).

Students With Learning Disabilities

Many students struggle as they try to put their thoughts onto paper or into a keyboard. However, written communication is often especially challenging for children and youth with LDs. Writing involves the adroit utilization and coordination of multiple processes: cognitive, linguistic, affective, behavioral, and physical (Santangelo, 2014), which can each present barriers to young people with LDs who "… fail to develop the knowledge, skill, will, and self-regulation necessary to succeed in key subject areas" (Grünke & Morrison-Cavendish, 2016, p. 1). For example, these students may have a less mature and developed understanding of the importance of writing. They may lack topic and genre knowledge, and may not know how to effectively engage in the various stages of the writing process. Thus, they often exhibit little to no structured or systematic planning before they commence the writing task (Gillespie & Graham, 2014). Instead, they engage in what Scardamalia and Bereiter (1986) termed "knowledge-telling" behavior, in which a person generates content by writing down whatever information he or she can recall about a topic without regard for the purpose or goal of the assignment. When revising, students with LDs tend to focus only on surface-level changes, such as punctuation and spelling, rather than evaluating their text to make meaningful modifications (Saddler & Asaro, 2007). Additionally, they may have difficulties with handwriting, spelling, and vocabulary, which may limit the amount of writing they produce and the words they choose to use. Finally, they may exhibit low motivation to write and low self-efficacy toward their writing ability (Graham, Collins, & Rigby-Willis, 2017).

Combined, these characteristics have a profound effect on the ability to produce quality written products. In a recent meta-analysis of the writing characteristics of students with LDs, Graham et al. (2017) found that, when compared with their typically achieving classmates, students with LDs produced writing samples that were of lower overall quality and less legible, more disorganized, and with more spelling and grammar errors and less diverse vocabulary and ideas.

Of all these areas of need, effectively organizing their writing may be one of the greatest challenges for students with LDs. Organization is a critical component of planning. During the planning process, writers compile relevant information and develop a blueprint that will assist them in accomplishing the goals of the composition. Along the way, they coordinate three subprocesses: generating (retriev-

ing relevant information), organizing (structuring the information), and goal setting (developing goals and establishing a writing plan to achieve the goals). This component of the overall writing process is so important that skilled writers spend considerable time planning and organizing. Further, it is the basis for the other two major text production activities – translating and reviewing (Flower & Hayes, 1981).

Unfortunately, children and youth with LDs minimize the role of planning and organizing (Scardamalia & Bereiter, 1986). They may have ideas about what they want to write, but they struggle to put their thoughts on paper and visually organize them (Sundeen, 2014). In fact, Koutsoftas (2016) found that many students with LDs may be able to produce as many as or more ideas than students without disabilities; however, they are less efficient at structuring and sequencing their thoughts. As a result, an organizational tool, such as a story map, may be beneficial for such students.

Story Mapping

A story map is a graphic organization tool that outlines the important components of a story. As such, story maps can support writers by providing visual reminders of these important elements while offering a space in which they can brainstorm and write down their notes for the respective element (Grünke & Leonard Zabel, 2015). Story maps also allow visualization of how the elements of a story – settings, characters, and events – are connected (Li, 2007). Such a tool is beneficial for writers with LDs, who tend to have difficulty generating stories (Barenbaum, Newcomer, & Nodine, 1987), and who write relatively short texts consisting of elements that are not logically related (MacArthur & Graham, 1987).

Story mapping has been found to be an effective intervention for students with LDs in both reading comprehension (Boon, Paal, Hintz, & Cornelius-Freyre, 2015) and writing (Hennes, Büyüknarci, Rietz, & Grünke, 2015; Li, 2007). Li (2007), for example, found that three of four fourth- and fifth-graders with LDs increased their writing fluency after learning a story-mapping strategy. This finding is supported by Hennes and colleagues (2015), who noted that when they received a story-mapping intervention, students between 8 and 14 years of age who had LDs significantly increased their number of words written, and, perhaps more important, the quality of their stories.

In two recent studies, peers were used to teach elementary school children with LDs how to apply story mapping to help them plan their narratives. The results indicate that the students with LDs significantly increased the number of words they wrote (Grünke, Janning, & Sperling, 2016; Grünke, Wilbert, Tsiriotakis, & Lopez Agirregoikoa, 2017). Both studies provide insight into the potential of using peer-assisted writing strategies to support students with LDs in story writing.

Peer-Assisted Writing

Peer-assisted writing strategies involve one student systematically assisting, or teaching, another in writing. Typically, this includes pairing a higher-achieving with a lower-achieving writer (Grünke et al., 2016). Peer-assisted writing offers several benefits to learners. First, peer support aids students by providing them with prompting, modeling, and immediate access to assistance that is individualized (Yarrow & Topping, 2001). In addition, peer tutors tend to limit processing overload that is often experienced by writers with LDs, allowing them to focus on higher-level processing rather than low-level skills such as spelling (Yarrow & Topping, 2001). Students who participate in peer writing structures also tend to demonstrate increased time on task and increased levels of engagement and experimentation (Englert, Berry, & Dunsmore, 2001; Yarrow & Topping, 2001). In a meta-analysis of writing interventions, peer-assisted writing was shown to have a positive impact on written products, with an overall effect size of 0.75, indicating a moderate to large effect (Graham & Perin, 2007).

A wide variety of students can benefit from peer assistance in writing, including students as young as kindergarten (Puranik, Patchan, Lemons, & Al Otaiba, 2017) to secondary students (Rensing, Vierbuchen, Hillenbrand, & Grünke, 2016), both with and without disabilities (e.g., Saddler & Asaro, 2008; Saddler & Graham, 2005), as well as English-as-a-foreign-language (EFL) students (Kurihara, 2017).

Research Question

The purpose of the current study was to contribute to the scarce body of empirical literature on the benefits of teaching story mapping to enhance the writing performance of students with LDs through peer tutoring. By testing an older population (three secondary school students with LDs), the study extends the two previously mentioned experiments by Grünke et al. (2016, 2017), which provide evidence that a peer-tutoring story-mapping strategy can help children at risk for failure to catch up with their classmates at a rather early stage of their writing development.

Method

Setting and Participants

The inclusive secondary modern school ("Hauptschule") chosen for this study is located in a metropolitan town in Northrhine-Westfalia (Germany). It enrolls approximately 240 students in grades 5 to 9. A "Hauptschule" offers lower secondary education (primarily for students with average grades and below) according to the International Standard Classification of Education (UNESCO, 2011). Any student who went to a four-year German elementary school can attend a "Hauptschule" afterwards, whereas attendance at a grammar school ("Gymnasium") – the other option for secondary education – requires high grades.

Three tutees and three tutors were chosen from one of the eighth-grade classes by the main teacher and the first author. Five of the students in this class had been diagnosed with LDs by a multidisciplinary team at the end of their elementary education. The process of identification had involved standardized assessment of language, reading, math, and nonverbal cognitive abilities. All five students had acquired basic skills in the aforementioned areas. However, they demonstrated profound difficulty applying them in higher-level schoolwork. Despite having an average IQ level, they had severe trouble applying problem-solving steps and transferring academic skills to other tasks, due to challenges in finding solutions to complex assignments and comprehending the logic behind them. There was a discrepancy of at least 1.5 standard deviations between aptitude (intelligence) and performance in higher-order skills (mathematical problem solving and reading comprehension).

We considered as potential participants of our study students who, according to their school records, did not score below the 30th percentile in a standardized spelling test. Thus, we wanted to make sure that participants were ready to attend to expressive writing without being held back by having to think too much about the correct order of letters in a word, for example. In addition, potential students had to be deemed by their main teacher as being socially capable of independently working in pairs without needing constant attention from an adult. Twelve students in the class met these criteria.

Prior to the start of the intervention, we presented these learners with a writing prompt consisting of a sequence of pictures in the form of a cartoon strip. We asked the students to imagine what was happening in those pictures and to write a story about it. No time limits were given. We allocated the three students who produced the shortest stories to function as tutees and assigned the three students with the longest texts as tutors. Pairs were built on the basis of the teacher's judgment of who got along best with whom. Not surprisingly, none of the five students with LDs were among the allocated tutors; however, all three tutees met the criteria for LDs.

The first team consisted of Aaron (tutee) and Adrian (tutor) (the participants' names have been changed to preserve their anonymity). Aaron was a 15-year-old male. According to his main teacher, he was comparatively difficult to motivate and demonstrated extremely weak analytical skills. Even though he possessed ample spelling abilities, he was a very reluctant writer. Adrian was 14 years old, and a student with average math and language skills. He enjoyed writing, and was viewed by his teacher as extraordinarily cooperative. Both boys were born in Germany and did not have an immigrant background.

The second team was made up of Baci (tutee) and Babak (tutor). Baci was 15 years old, and the daughter of Turkish migrant workers. Even though she was born in Germany, she spoke mostly Turkish at home with her family. Her teacher described her as having very low self-efficacy in writing and as not being performance-oriented. Baci's grades were generally in the D range. Babak was 14 years old, and the son of Iranian parents who migrated to Germany when he was a toddler. According to his teacher, he was very hardworking and organized. He demonstrated excellent written and verbal communication skills, and his grades were in the top quarter of his class.

The third and last team included two girls: Channa (tutee) and Cora (tutor). A daughter of Polish parents, 14-year-old Channa had moved with her family to Germany as an infant. Even though her spelling abilities were acceptable, her vocabulary was regarded as rather limited by her teacher. Her grades were in the range of Cs and Ds. She presented a notably low level of motivation and very little enthusiasm for learning. Cora was 15 years old, did not have an immigrant background, and was deemed by her teacher to be a very skilled reader and writer. She generally received grades in the A-B range, and was well known in her class for her prosocial behavior.

Experimental Design

We applied an AB multiple-baseline design across subjects (Gast, Lloyd, & Ledford, 2018). The

data were collected over 13 consecutive school days. Following the single-case reporting guidelines by Tate et al. (2016), a randomization procedure was used to increase the internal validity of the study, utilizing the randomize function in Microsoft® Excel. We stipulated that each phase in our design had to consist of at least three measurement points. Therefore, the start of the intervention for each of the three cases was chosen at random (with the restriction that there had to be at least three probes per participant in every phase). Hence, the treatment could have started any time between the 4th and the 11th probe. A random drawing of all six possible options for each participant resulted in an arrangement whereby the training for Aaron started after the sixth measurement point, and for Baci as well as for Channa after the eighth. Thus, Aaron received seven training sessions, and Baci and Channa each participated in five.

Materials

Baseline and intervention writing prompts consisted of fifty 5 x 8-inch index cards for each team that had story starters printed on one side (e.g., "One day, I discovered that I had a super power"). The prompts were based on the ideas in a book by Kinder (2014), which contains a high number of story starters of an equal difficulty level. Before printing them on the index cards, we simplified them

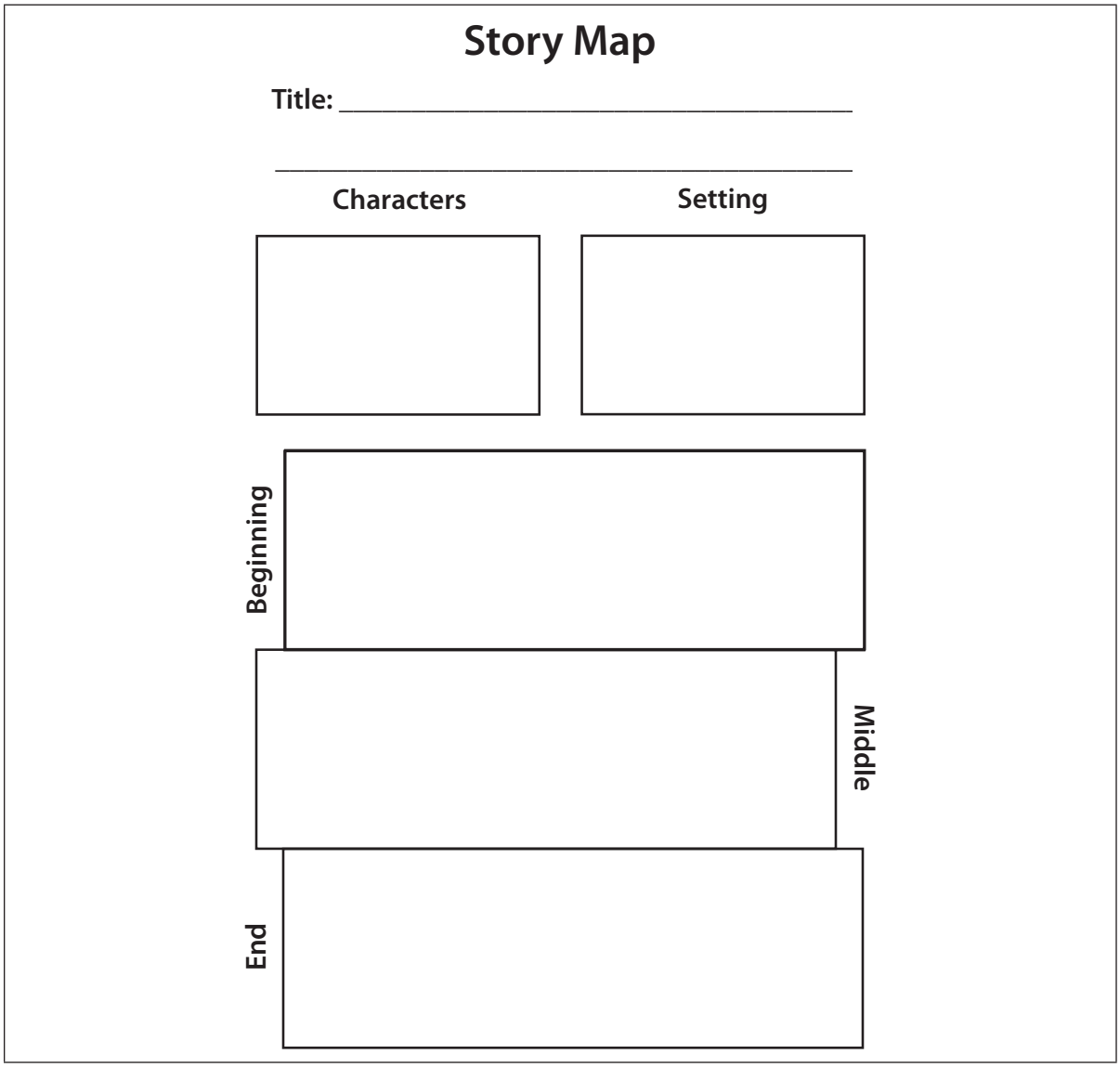

Figure 1. A story map template.

and translated them into German. At each of the 13 measuring points, the tutees were given a pen and four sheets of 8.5 x 11-inch notepaper.

For the intervention, we used a two-page list of 50 narrative writing prompts for sixth and seventh graders that were taken from the website https://k12.thoughtfullearning.com/resources/writingtopics (e.g., "If only I would have listened!,""Summer in a cabin by a lake," or "We couldn't stop laughing!"). Before presenting them to the participants, we translated the prompts into German. We employed simple story map templates as depicted in Figure 1. In addition, we provided each team with an 11.5 x 16.5-inch poster that visualized the steps of the mapping strategy: (a) contemplate 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 while using the appropriate fields. During the intervention, we made sure that the tutors and tutees always had enough pens and notepaper at their disposal. In order to guide the tutors through the lessons, and to provide them with reminders of what to do during the process, we prepared a simple four-page script containing brief instructions and mnemonics in large print (the scripts are available from the first author upon request).

Definition and Measurement of Dependent Variables

At each measurement point, a female research assistant asked the tutees to randomly draw two index cards with story starters from the pile. Aaron, Baci, and Channa were always given a choice to decide which of the two prompts they wanted to use as an initial point for their narratives. We never gave them a pile containing cards that they had already seen. The tutees were given a pen and notepaper. No time limits were set. The research assistant told the tutees that they should use the paper to produce their stories, but that they could also utilize some of it to outline their ideas if they felt a need to do so.

Two dependent variables were used to capture the writing performance of Aaron, Baci, and Channa: the number of total words written (TWW; Hosp, Hosp, & Howell, 2016) and a writing rubric developed by Harris and Graham (1996). The TWW is a widely used production-dependent fluency measure (Furey, Marcotte, Hintze, & Shackett, 2016) defined as the number of recognizable words written, regardless of spelling or context (excluding digits). Any letter or group of letters that has a space before and after it (even if it must be viewed as a nonsense word) is considered a word (Hosp et al., 2016). The writing rubric was included to determine the completeness of the texts. It consists of eight categories of descriptors: main character, locale, time, starter event, goal, action, ending, and reaction (see Figure 2). The rubric includes specifications for how certain criteria must be met in order to earn a specific number of points for one of the eight criteria. In total, a text could be awarded between 0 and 19 points.

Two female graduate students, who were blind to the purpose of the study, served as scorers. Both of them had received extensive training by the first author on how to use the instruments. First, they determined the total number of words the tutees had written (TWW). They randomly chose one text after another and evaluated them. Subsequently, they counted TWW independently and compared their results. Inter-rater agreement was calculated for each text by dividing the smaller number by the larger and multiplying by 100. The scores varied between 95.15 and 100.00%, with an average of 97.05%. In case of discrepancies, the raters discussed them until they reached consensus.

Second, one of the graduate students scored all the stories using the writing rubric whereas the other independently scored a random selection of 20% of the texts. Inter-rater agreement was determined using the same procedure as with TWW. Agreement ranged between 73.33 and 100.00%, with an average of 88.48%. Such a value meets the standards established by Hartmann, Barrios, and Wood (2004) for percentage agreements in single-case designs. We used the scores obtained by the first graduate student for the data analysis. The second rater was involved to gain an indication of the reliability of the scores.

Procedures

Baseline. Each day of the study, at the beginning of the third class period, the research assistant picked up the tutor-tutee teams from their class and brought them to a resource room in the school. The rest of the students in the class remained with their main teacher and engaged in independent reading or writing activities.

In the resource room, the teams were seated at tables as far apart as possible. For about 30 minutes, they participated in different educational games (such as math racetracks; see Skarr et al., 2014). (None of these activities involved story writing or similar tasks.) Then the research assistant gave Adrian, Babak, and Cora some math problems to solve, whereas Aaron,

A Scale for Scoring the Inclusion and Quality of the Parts of a Story

1. Main Character

0	No main character is established.
1	A main character is presented; however, he/she is just a name on a page. Very little information or detail about the main character is provided.
2	A main character is presented and described in such detail that he/she is always "real" for you.

2. Locale

0	No locale or place is mentioned.
1	Locale given, but little description offered.
2	Locale given, with more complete description offered, or unusual locale is chosen.

3. Time

0	No time given.
1	Time given, but traditional in reference.
2	Time given, but unusual in reference or more complete description.

4. Starter Event

0	The precipitating event that causes the main character to establish a goal is not presented.
1	The precipitating event that causes the main character to establish a goal is presented. The precipitating event can be a natural occurrence, an internal response, or an external action.
2	The precipitating event is complex, unusual, or well described.

5. Goal

0	The goal or purpose of the main character is not established.
1	The goal or purpose of the main character is established but not clearly articulated.
2	The goal or purpose of the main character is clearly articulated.
3	Two or more goals are clearly articulated.

6. Action

0	The actions that the main character initiates in order to achieve the goal are not presented.
1	What the main character does in order to achieve the main goals is presented.
0-4	Add one point for each of the following: A. Actions or events happen in a logical order (i.e., they are not inconsistent). B. Ingenuity or originality are used to solve situations or predicaments. C. There is more than one well-defined episode. For example, the main character tries one action, and if it is unsuccessful then tries another action. D. The main character goes to one place during his travels and then another.

7. Ending

0	No real ending, lack of conclusion, or story seems unfinished. In other words, the long-range consequences of the main character's actions are not resolved.
1	Long-range consequences of main character's actions are resolved, but the ending or conclusion is fairly common.
2	Long-range consequences of main character's actions are resolved. In addition, the conclusion or ending is unusual, or the ending contains a moral.

8. Reaction (expressed anywhere in the story)

0	The emotional reactions of the main character are not presented.
1	Some emotional feelings expressed by the main character.
2	Emotions feelings of the main character expressed with depth.

Figure 2. The writing rubric by Harris and Graham (1996) (the original version provides examples for the different categories).

Baci, and Channa were asked to write a story in accordance with the description in the above section on measurement of dependent variables.

Tutor preparation. Before the teams started with the story-mapping intervention, the research assistant provided the tutors with a two-hour training on the components of the instructional framework, including modeling and strategy instruction. She also familiarized them with the script and encouraged them to refer to it as a memory aid.

Intervention. In the treatment condition, the measuring procedures were identical to those of baseline. However, the teams no longer engaged in educational games, but practiced outlining acceptable narratives using the story-mapping strategy. Each lesson lasted 30 minutes. As in the baseline condition, the research assistant was always present to provide support if needed. Training sessions followed the common scaffolding sequence for direct instruction of "I do it, we do it, you do it" (Archer & Hughes, 2010). The tutees often had trouble concentrating during regular school lessons, but being in a situation where they received individual instruction helped them to stay on task. Even the fact that there were three teams in the same room, working on different tasks, hardly seemed to bother or distract anyone. However, if a team got sidetracked or needed additional support, the research assistant intervened.

Lesson 1. In the first lesson, the tutors put the poster with the strategy steps on the table. They read each step out loud and assured their partners that using this technique would help them to write better stories. The tutors then reverted to the two-page list of narrative writing prompts. After choosing one of the headings, they crossed it out and modeled the process by performing the activities listed on the poster using completed story-map templates while thinking aloud. The tutors then repeated the steps of the strategy and asked the tutees follow-up questions to make sure that they understood the purpose of the procedure.

In the remaining time, Aaron, Baci, and Channa filled out their own story maps, while Adrian, Babak, and Cora provided scaffolded feedback and encouragement. Before the tutees' performance was measured, the tutors took the posters away in order to create a situation that conformed with the baseline conditions.

Lesson 2. At the beginning of the second lesson, the tutors again placed the posters on the table and went over the steps of the strategy once more to help the tutees to internalize them. Afterwards, tutors and tutees together selected a story starter from the list, crossed it out, and took turns filling out the fields of the template while executing the four activities on the poster. In the remaining minutes, they repeated the process at least one more time. Finally, the posters were removed, and the tutees wrote a story to measure their performance.

Lesson 3 and remaining lessons. From the third session onward, support from the tutors was gradually reduced to provide the tutees with opportunities for independent practice. The posters were again placed on the table, but this time they were not explicitly mentioned; they just served as a memory aid.

Aaron, Baci, and Channa were now encouraged to choose a writing prompt from the list and fill out a template by themselves. If necessary, Adrian, Babak, and Cora gave corrective feedback, and if the process came to a halt, they offered suggestions on what to fill into the fields. At the end of the lesson, the tutees were prompted to create their own story maps without using one of the pre-prepared templates. After about half an hour, the tutees' performance was measured.

During the remaining sessions, the tutees created their own story maps and were constantly encouraged by their tutors as they devised ideas to enter into the fields of the template. Aaron, Baci, and Channa did not take notes during performance measurements under baseline conditions, but they started doing so after the second treatment session. Moreover, they quickly began to create their own story maps in order to organize their ideas prior to the writing process.

Treatment Fidelity

The research assistant used a checklist to ensure fidelity of implementation. It contained all relevant aspects of the procedure as described above. If the tutors deviated from the script, the research assistant interceded and made sure that the treatment was executed the way in which it was initially planned.

Social Validity

After the completion of the last training session, the research assistant interviewed Aaron, Baci, and Channa individually, asking their opinion about the peer-tutoring intervention. Specifically, she asked (a) if they liked the format of the treatment, (b) if they viewed it as helpful, (c) if they thought it had an impact on their writing performance, and (d) if they would recommend it to other students. The research assistant took notes on the participants' feedback.

Results

Descriptive Analysis

Tables 1 and 2 present the descriptive statistics for the TWW and the writing rubric scores. As illustrated, the two dependent variables are highly correlated across cases ($r_{Spearman's\ Rank} = .85$; $p < .001$; one-tailed). In examining the raw scores, it is noteworthy that Channa wrote an exceptionally long story of 359 words after the fourth treatment session. The rest of the narratives that she produced during Phase B were, on average, not even half as long. She may have had an especially creative day and, therefore, been able to demonstrate her capabilities. A comparison between the mean TWW values during baseline and the intervention revealed an increase in performance for Aaron of 112.24%; Baci, 183.17%; and Channa, 49.17%. The rubric rating scores grew by 75.70% for Aaron; Baci, 115.38%; and Channa, 121.09%. In sum, these treatment gains are remarkable.

Visual Analysis

A visual display of the data is provided in Figures 3 and 4 (the graphs were produced using the MultiSCED web application by Cools et al., 2018).

Table 1
The TWW for Each Participant

		Baseline	Intervention
Aaron	*N* (Probes)	6	7
	Raw Scores	81; 76; 70; 88; 47; 79	164; 131; 208; 137; 156; 152; 188
	M	73.50	156.00
	SD	14.27	27.44
	Range	47-88	131-208
Baci	*N* (Probes)	8	5
	Raw Scores	88; 68; 79; 68; 110; 59; 66; 92	176; 223; 192; 318; 307
	M	78.75	223.00
	SD	16.99	65.59
	Range	59-110	176-318
Channa	*N* (Probes)	8	5
	Raw Scores	79; 130; 93; 90; 109; 92; 60; 71	106; 135; 129; 359; 192
	M	90.50	135.00
	SD	21.89	102.71
	Range	60-130	106-359

Table 2
Writing Rubric Scores for Each Participant

		Baseline	Intervention
Aaron	*N* (Probes)	6	7
	Raw Scores	7; 7; 9; 7; 5; 6	12; 11; 12; 12; 13; 12; 12
	M	6.83	12.00
	SD	1.33	0.58
	Range	5-9	11-13
Baci	*N* (Probes)	8	5
	Raw Scores	6; 6; 6, 5, 10; 6; 4; 9	13; 14; 15; 13; 15
	M	6.50	14.00
	SD	2.00	1.00
	Range	4-9	13-15
Channa	*N* (Probes)	8	5
	Raw Scores	5; 4; 6; 8; 7; 5; 4; 8	11; 11; 13; 14; 13
	M	5.88	13.00
	SD	1.64	1.34
	Range	4-8	11-14

For both the TWW and the writing rubric as the outcome score, the visual analysis for the within-condition comparisons in the baseline settings demonstrates a lack of trend, little variability, and shows that the level represented the data points well. In the treatment sessions, the within-condition comparisons suggest an increasing trend in the desired direction. The between-condition comparisons, in turn, reveal a large difference in level, in that higher TWW scores were observed in the sessions preceded by the intervention compared with the no-intervention sessions. An experimental effect can also be determined by the proportion of overlap between the data in different conditions, where fewer overlapping data points indicate a relatively stronger effect (Kratochwill et al., 2010).

We next applied an non-overlap index, the Percentage of Non-Overlapping Data (PND) (Scruggs, Mastropieri, & Casto, 1987), which is defined as the ratio of measurements in Phase B that exceeds the highest measurement from Phase A. For TWW, the PND for Aaron, Baci, and Channa equaled 100%, 100%, and 60%, respectively. For the writing rubric, the PND equaled 100% across the three participants. According to Scruggs and Mastropieri (1998), PND scores above 90 are indications of very effective treatments, scores from 70 to 90 represent effective interventions, scores from 50 to 70 represent

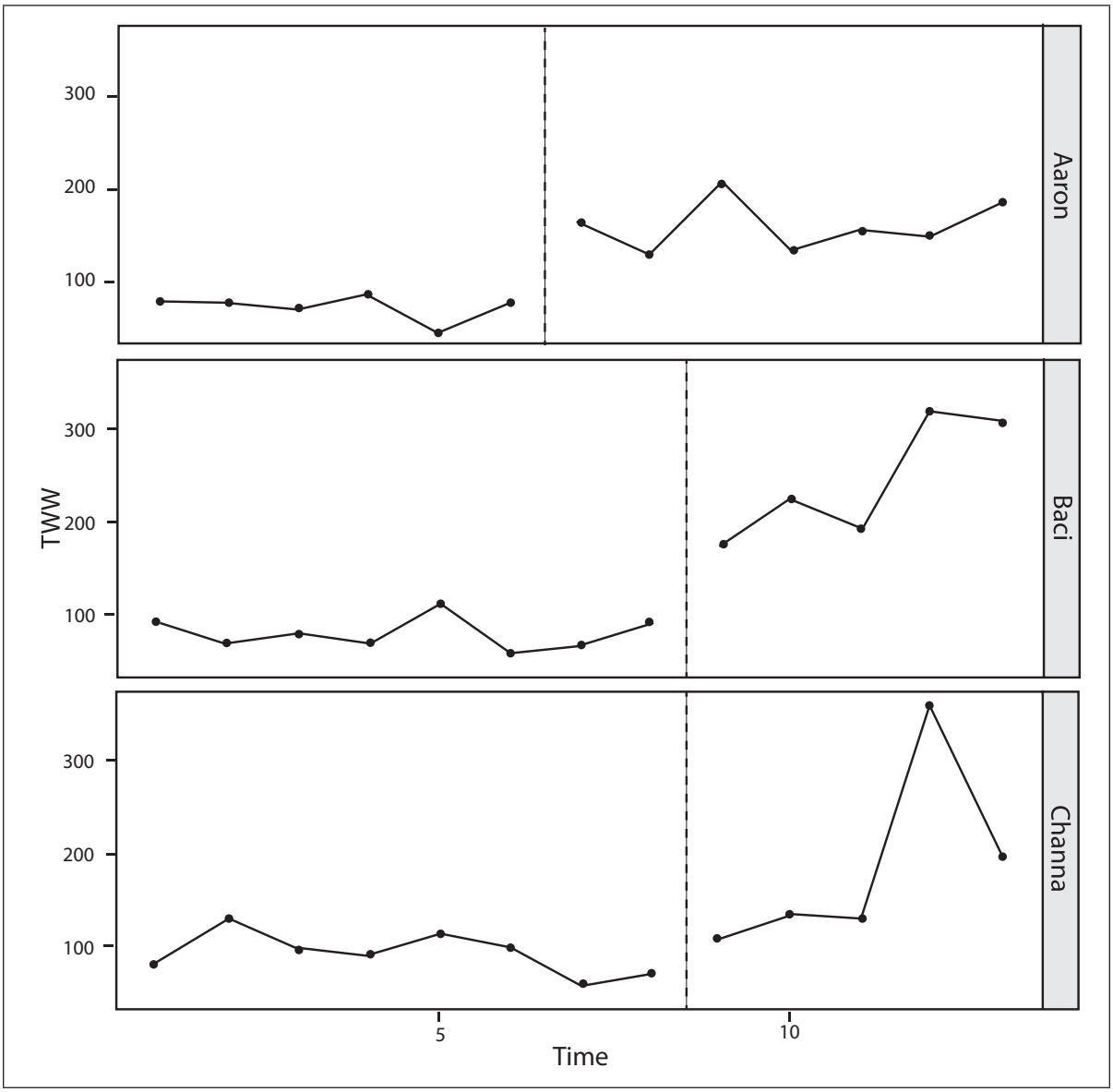

Figure 3. TWW for all three cases.

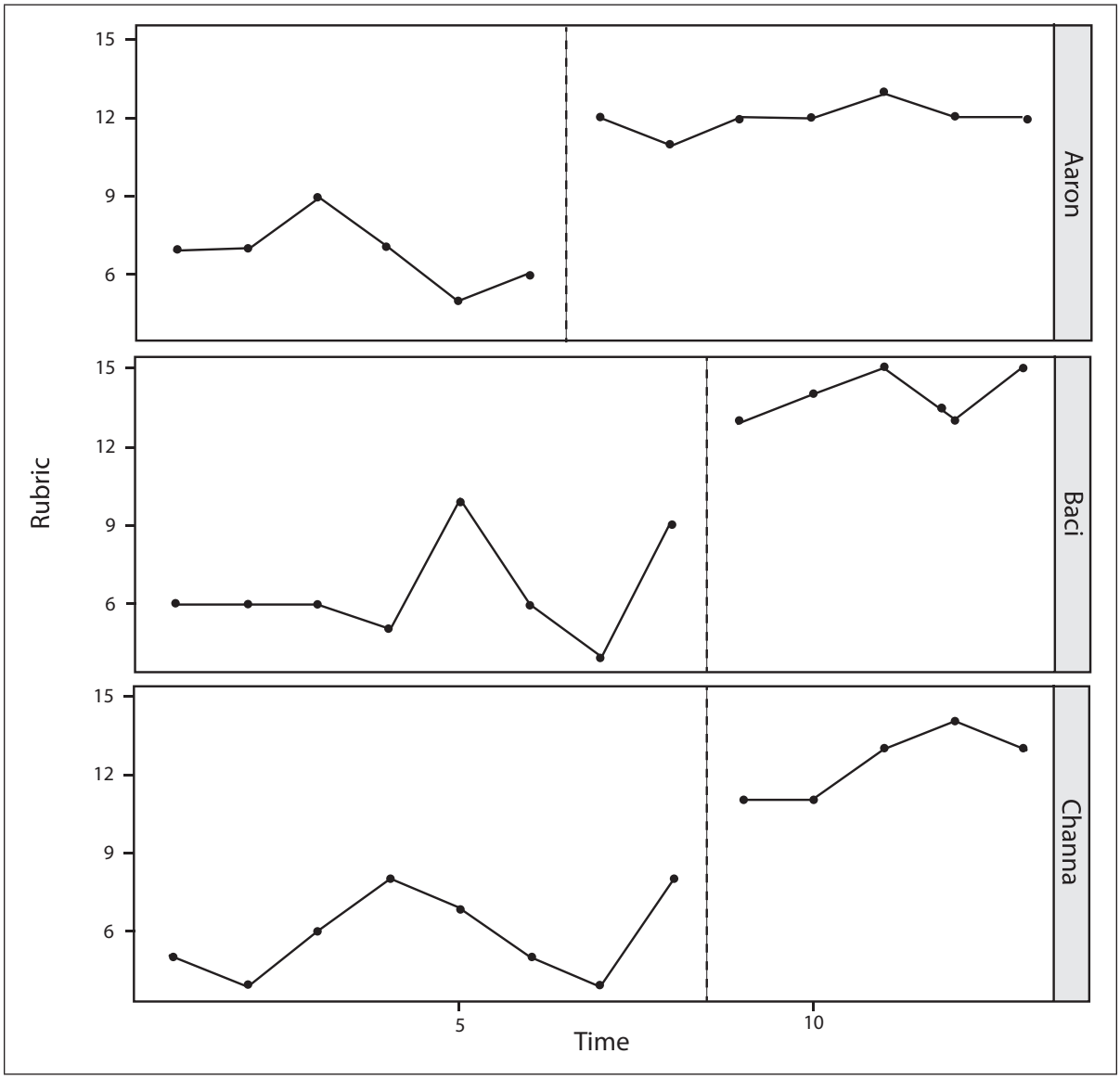

Figure 4. Writing rubric scores for all three cases.

questionable interventions, and scores below 50 are indicative of ineffective interventions. Hence, overall, the strategy instruction may be considered as being very beneficial. Only the PND score of 60% for Channa in the TWW stands out. As mentioned earlier, the number of words that she wrote in her stories increased from Phase A to Phase B by about 50%. However, she produced one relatively short text of just 106 words at the beginning of her relatively brief intervention, which strongly contributed to her low PND score.

Quantitative Analysis

Horner et al. (2005) consider visual inspection the traditional and most common mode of analyzing data from single-case research. However, increasingly more researchers, including the authors of this paper, are calling for a complementary use of statistical techniques to provide a direct test of the null hypothesis and to employ precisely defined criteria for significance (Grünke, Boon, & Burke, 2015; Lobo, Moeyaert, Baraldi, & Babik, 2017; Stone, Friedlander, & Moeyaert, 2018; Tate et al., 2016). As a result, we also performed the following tests.

Randomization test. As previously mentioned, treatment for each of the three cases started at random with the only restriction being that there should be at least three measurements per participant in each phase. Therefore, the total number of unique ways in which the start of the intervention could be delivered equals $8^3 = 343$. The *p* value assesses the null hypothesis that the experimental manipulation had no impact on the dependent variable by "rearranging the observed scores to all permutations of the possible randomization orders and examining different outcomes" (Barlow, Nock, & Hersen, 2009, p. 284).

In the present study, we conducted the randomization test using the SCDA package within R (Bulté & Onghena, 2013). The difference between the mean of the treatment phase scores and the mean of the baseline phase scores on each dependent variable was used as test statistic (i.e., $\bar{B} - \bar{A}$; Edgington & Onghena, 2007) and compared with a sampling distribution using the set of 343 permutations. For the TWW, the observed test statistic was 115.64 ($p < .01$); for the writing rubric, it was 6.40 ($p < .01$). As a consequence, we can deduce that the story-mapping strategy intervention had a statistically significant and large effect on both text quantity and completeness.

Individual-level analysis. The following data evaluation focused on calculating different effect size measures to determine the quantitative magnitude of the treatment benefits. For our single-level analysis, we included Glass' Δ (Glass, McGaw, & Smith, 1981), Hedges' *g* (Hedges, Pustejovsky, & Shadish, 2013), the Tau-U (Parker, Vannest, Davis, & Sauber, 2010), and a regression-based approach (Van den Noortgate & Onghena, 2003a, 2003b). The advantage of the regression-based estimate is that, in addition to the changes in level, it is possible to estimate the initial baseline level, the trend during the baseline, the immediate treatment effect, and the change in trend between the baseline and treatment phase (Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, 2014).

From the single-level analysis, we can deduce that the intervention was consistently statistically significant across a variety of effect sizes for both the TWW and the writing rubric as the outcome (see Table 3). This finding offers evidence that the story-mapping strategy is an effective intervention.

Two-level analysis. In addition to calculating the Tau-U for each student, we generated a weighted, across-case Tau-U, using the Tau-U online calculator (Vannest, Parker, & Gonen, 2011). The analysis yielded a weighted Tau-U of .95 [$SD = 0.20$; $z = 4.86$; $p < .01$) for the TWW and 1.00 [$SD = 0.20$; $z = 5.11$; $p < .01$) for the writing rubric. Thus, the treatment effect can be considered to be remarkable.

Moreover, after ensuring that the data met the requirements for conducting a regression analysis, we applied multilevel modeling, as outlined by Van den Noortgate and Onghena (2008). The purpose of this last step of the evaluation was to estimate the overall average effect of the treatment across participants, as well as the within- and between-case variability.

Two different two-level models were run. Equations 1 through 3 present "Model 1," in which the overall average baseline level and the overall average change in level is estimated in addition to the between-case variance in the baseline level and between-case variance in

Table 3
Output Effect Sizes for the Single-Level Analysis

		Effect size type			
Variable	Case	Glass' Δ	Hedges, *g***	Tau-U	Regression****
TWW	Aaron	5.78	3.68	1.00* [*SD* = 0.33]	88.79 [*SE* = 12.48]*
	Baci	8.49	3.65	1.00* [*SD* = 0.34]	164.45 [SE = 23.84]*
	Channa	2.03	1.35	0.85* [*SD* = 0.34]	93.70 [SE = 36.69]*
Rubric	Aaron	3.89	4.84	1.00* [*SD* = 0.33]	5.17 [*SE* = 0.55]*
	Baci	3.75	4.09	1.00* [*SD* = 0.34]	7.50 [*SE* = 0.97]*
	Channa	6.93	3.94	1.00* [*SD* = 0.34]	6.53 [*SE* = 0.88]*

* $p < .001$.
**Hedges bias-corrected effect size for small samples.
***Changes in level effect size.

changes in level. In Equation 1, y_{ij} refers to the outcome at measurement occasion i nested within participant j. The outcome is regressed on an intercept and a dummy-coded variable ($Treatment_{ij}$) indicating to which phase the outcome score belongs. If $Treatment_{ij}$ is 0, then the score is part of the baseline phase; otherwise it is part of the treatment phase. Therefore, β_{0j} is the parameter reflecting the baseline level for case j, and β_{1j} represents the change in level, hence the treatment effect for case j.

Level 1:

$$y_{ij} = \beta_{0j} + \beta_{1j} Treatment_{ij} + e_i \text{ with } e_i \sim N(0, \sigma_e^2) \quad (1)$$

It is unlikely that the baseline level and the treatment effect is the same for all participants. Therefore, a second level is added to the model, allowing the subject-specific coefficients (and) to vary across subjects:

Level 2:

$$\begin{cases} \beta_{0j} = \theta_{00} + u_{0j} \\ \beta_{1j} = \theta_{10} + u_{1j} \end{cases} \text{ with } \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{u_0} & \\ \sigma_{u_1 u_0} & \sigma^2_{u_1} \end{bmatrix} \right) \quad (2)$$

Using equations 1 and 2, the overall average baseline level (θ_{00}) and intervention effect (θ_{10}) can be estimated in addition to the between-case variability in baseline level ($\sigma^2_{u_0}$) and treatment effectiveness ($\sigma^2_{u_1}$).

Equations 5 and 6 present "Model 2;" they were used to estimate the overall average baseline level (θ_{00}), trend during baseline (θ_{10}), immediate treatment effect (θ_{20}) and treatment effect on the time trend (θ_{30}), in addition to the between-case variance in these estimates.

Table 4
The Output Effect Sizes for the Two-Level Analysis

		Model 1			*Model 2*		
Variable		Estimate	*SE*	*p*	Estimate	*SE*	*p*
TWW	**Fixed Effects**						
	Baseline level	81.59	9.63	<.001	70.29	17.07	< .001
	Trend baseline	/	/	/	-2.67	3.58	.46
	Immediate treatment effect	114.19	26.01	.028	76.25	25.62	.021
	Treatment effect on trend	/	/	/	26.86	12.74	.126
	Variance Effects						
	Baseline level	0.00	/	/	0.00	/	/
	Trend baseline	/	/	/	0.00	/	/
	Immediate treatment effect	1384.89	1753.36	.21	353.28	1031.25	.366
	Treatment effect on trend	/	/	/	347.12	394.43	.189
	Residual variance	2040.28	487.70	<.001	1359.64	343.83	< .001
Rubric	**Fixed Effects**						
	Baseline level	6.36	0.30	<.001	6.76	0.65	< .001
	Trend baseline	/	/	/	0.08	0.15	.61
	Immediate treatment effect	6.41	0.68	.002	5.40	1.01	< .001
	Treatment effect on trend	/	/	/	0.20	0.25	.43
	Variance Effects						
	Baseline level	0.00	/	/	0.00	/	/
	Trend baseline				0.012	0.03	.32
	Immediate treatment effect	0.74	1.12	.25	0.84	1.25	.25
	Treatment effect on trend				0.00	/	/
	Residual variance	2.07	0.49	<.001	1.93	0.49	< .001

Level 1:

$$y_{ij} = \beta_{0j} + \beta_{1j}Time_{ij} + \beta_{2j}Treatment_{ij} + \beta_{3j}Treatment_{ij} * Time_{ij} + e_i \text{ with } e_i \sim N(0, \sigma_e^2) \quad (4)$$

Level 2:

$$\begin{cases} \beta_{0j} = \theta_{00} + u_{0j} \\ \beta_{1j} = \theta_{10} + u_{1j} \\ \beta_{2j} = \theta_{20} + u_{2j} \\ \beta_{3j} = \theta_{30} + u_{3j} \end{cases} \text{ with } \begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \\ u_{3jk} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{u_0}^2 & & & \\ \sigma_{u_1u_0} & \sigma_{u_1}^2 & & \\ \sigma_{u_2u_0} & \sigma_{u_2u_1} & \sigma_{u_2}^2 & \\ \sigma_{u_3u_0} & \sigma_{u_3u_1} & \sigma_{u_3u_2} & \sigma_{u_3}^2 \end{bmatrix} \right) \quad (5)(6)$$

For more details and interpretation of the two-level analysis of single-case experimental designs, we refer the reader to Moeyaert, Ferron, Beretvas, and Van den Noortgate (2014). The output for the present study may be found in Table 4.

From the two-level analysis, we can deduce that, across the three cases and the two models, the treatment appears to be statistically significant at the .05 level (two-tailed). For Model 1, the change in the average TWW between the baseline and treatment phases equaled 114.19 [$t(2.68) = 4.39$, $p = .028$]. For Model 2, the immediate treatment effect (the change in the TWW between the end of the baseline phase and the start of the treatment phase) equaled 25.62 [$t(2.68) = 4.39$, $p = .028$]. For Model 1, the change in the average writing rubric between baseline and treatment phase equaled 6.41 [$t(3.09) = 9.41$, $p = .002$]. For Model 2, the immediate treatment effect (the change in the writing rubric between the end of the baseline phase and the start of the treatment phase) equaled 5.40 [$t(10.4) = 5.32$, $p < .001$].

Qualitative Analysis

Responses from the interviews conducted by the research assistant, which lasted 8-10 minutes each, helped to verify that all three tutees considered the intervention to be very beneficial. However, it needs to be taken into consideration that the survey was conducted in a rather informal way. That is, we did not choose a particular method from the pool of elaborated and well-established approaches in this field of research. Rather, after the intervention, the research assistant took each of the participants aside to ask them if they liked the format of the treatment, if they viewed it as helpful, if they thought it had an impact on their writing performance, and if they would recommend it to other students, to get some general feedback on how the treatment was received by the students (see Snodgrass, Chung, Meadan, & Halle, 2018).

The responses of the three tutees suggest that the peer-tutoring instruction was well received and that they enjoyed participating in the study. Indeed, they all expressed their approval of the treatment very emphatically. Baci stated: "I normally don't like writing; I like reading better. But working together with Babak was fun." All three tutees also stressed that they deemed the intervention as extremely helpful as they tried to structure their ideas before actually writing their stories. Aaron commented: "Finally, someone showed me how to write better. Story maps make everything much easier." The three children appeared to be very proud of their accomplishments and were under the impression that their writing had improved significantly. They attributed the increases in performance to the format of the instruction. Finally, they unanimously stated that they would recommend the treatment to other children who are struggling with text composition. Channa explained: "It feels good to be able to write much longer stories now. I believe that other kids would also find story maps helpful."

Discussion

Main Findings

The purpose of this single-case analysis was to determine the effects of a peer-tutoring graphic organizing strategy on the writing performance of three eighth graders with LDs. Few studies have evaluated the benefits of teaching text-planning skills to struggling learners by using fellow students as instructors. In this experiment, we demonstrated that a rather short intervention of five to seven 30-minute lessons resulted in statistically and practically significant improvements in the length and completeness of narratives written by the tutees.

The performance enhancements reached impressive magnitudes ranging from around 50% to over 180% in the TWW and from about 75% to a little over 120% in the writing rubric. In five of six instances, there was no overlap between the data points during the baseline and intervention. A randomization test revealed a significant overall level change between phases. On an individual basis, different effect sizes indicated remarkable treatment outcomes for all three tutees. Finally, our summary of the data across the three cases via an overall Tau-U and multilevel modeling confirmed the results from all previous analyses and highlight the great benefit that the peer-tutoring intervention had on the performance of the participants.

Thus, the results are in line with our expectations and with the findings from extant studies on teaching students with LDs the story-mapping strategy

through peer tutoring (see Grünke et al., 2016, 2017). What is especially gratifying and unique with regard to our experiment is that it shows that, although early intervention is important, it is not too late to teach adolescents with LDs simple text-planning skills using rather simple means. That is, it did not take long for the students to be able to apply the strategy and significantly improve their text-production skills. According to their main teacher, they had been trying unsuccessfully for many years to become more proficient writers. Furthermore, they seemingly enjoyed the intervention and were very happy about the progress that they had made.

Limitations

Notwithstanding the promising results, the study is subject to certain limitations. First and foremost, the small number of participants does not allow for generalization of the data. However, according to the widely accepted standards for single-subject research by Horner et al. (2005), external validity can be achieved through replication. The Council for Exceptional Children (CEC; 2014) stipulates that at least five methodologically sound case reports with positive effects and at least 20 total participants are needed in order to consider a treatment as evidence based. Thus, this experiment is a step toward the goal of establishing the intervention as an approach that meets these standards.

We incorporated a randomization procedure into the design, which increased the internal validity of the single-case analysis. Fortunately, we ended up with at least six baseline measurements, thus exceeding the minimum standard by Horner et al. (2005). The data trend prior to the intervention was quite stable, which also speaks to a high internal validity. However, the final two baseline data points for all three tutees indicate a possible rise in performance rather than level or decreasing performance. To strengthen the assumption of a functional connection between treatment and achievement, additional baseline measurements would have been beneficial. In addition, we did not provide for follow-up data collection. Even though the participants learned quickly how to make efficient use of the strategy, there is no way to know – with a given degree of certainty – whether the intervention effects are lasting. Upcoming school holidays did not allow for the acquisition of additional data directly upon the termination of the treatment. Undoubtedly, information about the stability of the performance gains would have helped to make an even stronger case for applying our intervention in schools.

Furthermore, the process of selecting participants may be considered a limitation. To measure the tutees' learning progress, we asked them to compose texts in response to story starters. However, we predicated the identification of eligible participants on the number of words the students produced while writing about a cartoon strip. Even though the fact that the TWW of the three tutees averaged less than 100 and the number of total points in the rubric averaged less than 7 during baseline speaks to the supposition that our identification process fulfilled its intended purpose, a different option for choosing the participants might have been more suitable. For example, we could have utilized writing prompts with no pictures to better align the selection procedure with the intervention. In addition, we could have considered the quality of the texts while trying to find tutees who would especially benefit from the treatment.

A final limitation pertains to the way the dependent variable was measured. While capturing TWW is the most common way to monitor writing performance, it is not the only one (see, e.g., Dockrell, Connelly, Walter, & Critten, 2015; McMaster & Espin, 2007). The same goes for the writing rubric that we used to measure students' skills in producing narratives. As writing is a complex ability, no one way of measuring could ever capture it completely. Hence, the way in which we performed our assessments is open to criticism. However, although we approached writing skills from two rather different angles (TWW and a writing rubric), the results were highly correlated. Moreover, both procedures were clearly very sensitive to change and were able to reflect the improvements that the participants had undergone.

Implications for Practice and Future Research

Despite any weaknesses that this study might have, the present findings are very encouraging and suggest that despite being such a demanding task, writing can be effectively taught to youth with LDs with very minimal resources. The intervention consisted of only five to seven 30-minute lessons, and instruction for the tutors only took two hours. The treatment was also inexpensive, including the cost of printing the posters, the story-map templates, and the index cards. Finally, ways were found for the teams to work on their assignments, even though the school was not able to provide an individual room for each tutor-tutee pair.

As Johnson and Semmelroth (2014) rightly pointed out, "while arguably no other content area in education has produced more instructional practice research than special education, the profession itself has made little progress in getting these instructional strategies into practice" (p. 71). One major reason for this gap stems from the fact that many teachers are overly burdened with crowded classrooms full of very diverse learners and a wide range of administrational tasks unrelated to teaching (Vannest & Hagan-Burke, 2009). This is particularly true in special education settings (Greenglass, Burke, & Konarski, 1997).

Being able to impart a highly effective strategy such as story mapping by way of peer tutoring opens up opportunities for teachers to provide students with LDs and other struggling learners with much-needed individually tailored support. In this study, the treatment was monitored by a research assistant. Even though the main teachers should generally be capable of overseeing peer-tutoring interventions in their own classrooms, an additional person to provide help whenever necessary is beneficial. To that end, Christle and Schuster (2003) suggested that educators involve interns, administrators, college students, school aides, volunteers, and other temporarily available people into everyday teaching.

Future research is warranted to confirm our findings with larger samples and to investigate the duration of the effects of an intervention such as ours. Such experiments should be conducted using different empirical designs and different ways of measuring text production abilities. In addition, it appears meaningful to not only monitor the tutees' progress, but also to determine the impact on the tutors of a system of instruction in which learners help each other. If tutees also benefited, it would constitute an even stronger case for implementing the strategy into everyday life in school than if only the tutees profited. Finally, the focus in this study was on improving the planning skills of adolescents with LDs in order to help them to write better narratives. Future research should also zero in on supporting these students to produce other text genres (e.g., descriptions, expositions, or poetry) and to foster other writing subskills (e.g., composing, refining, editing, or reviewing).

References

Applebee, A. N., & Langer, J. A. (2011). A snapshot of writing instruction in middle schools and high schools. *English Journal, 100,* 14-27.

Archer, A. L., & Hughes, C. A. (2010). *Explicit instruction: Effective and efficient teaching*. New York, NY: Guilford.

Barenbaum, E., Newcomer, P., & Nodine, B. (1987). Children's ability to write stories as a function of variation in task, age, and developmental level. *Learning Disability Quarterly, 10*, 175-188.

Barlow, D. H., Nock, M., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior for change*. London, UK: Pearson.

Boon, R. T., Paal, M., Hintz, A. M., & Cornelius-Freyre, M. (2015). A review of story mapping instruction for secondary students with LD. *Learning Disabilities: A Contemporary Journal, 13*, 117-140.

Bulté, I., & Onghena, P. (2013). The single-case data analysis package: Analysing single-case experiments with R software. *Journal of Modern Applied Statistical Methods, 12*, 450-478. https://doi.org/10.22237/jmasm/1383280020

Council for Exceptional Children (CEC). (2014). Standards for evidence-based practices in special education. *Teaching Exceptional Children, 46*, 206-212.

Christle, C. A., & Schuster, J. W. (2003). The effects of using response cards on student participation, academic achievement, and on-task behavior during whole-class math instruction. *Journal of Behavioral Intervention, 12*, 147-165.

Cools, W., Declerq, L., Beretvas, S. N., Moeyaert, M., Ferron, J. M., & Van den Noorgate, W. (2018). *MultiSCED user guide*. https://kuleuven.app.box.com/v/MultiSCEDUserGuide

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.

Edgington, E. S., & Onghena, P. (2007). *Randomization tests*. Boca Raton, FL: Chapman & Hall/CRC.

Englert, C., Berry, R., & Dunsmore, K. (2001). A case study of the apprenticeship process. *Journal of Learning Disabilities, 34*, 152-171. https://doi.org/10.1177/002221940103400205

Flower, L. S., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*, 365-387.

Furey, W. M., Marcotte, A. M., Hintze, J., & Shackett, C. (2016). Concurrent validity and classification accuracy of curriculum-based measurement for written expression. *School Psychology Quarterly, 31*, 369-382. https://doi.org/10.1037/spq0000138

Gast, D. L., Lloyd, B. P., & Ledford, J. R. (2018). Multiple baseline and multiple probe designs. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (pp. 239-282). New York, NY: Routledge.

Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children, 80*, 454-473. https://doi.org/10.1177/0014402914527238

Glass, G.V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. London, UK; Sage.

Graham, S. (2008). *Effective writing instruction for all students*. Wisconsin Rapids, WI: Renaissance Learning.

Graham, S., & Hall, T. E. (2016). Writing and writing difficulties from primary grades to college: Introduction to the special issue. *Learning Disability Quarterly, 39*, 3-4. https://doi.org/10.1177/0731948715592154

Graham, S., Capizzi, A., Harris, K. R., Herbert, M., & Morphy, P. (2014). Teaching writing to middle school students: A national survey. *Reading and Writing, 27*, 1015-1042. https://doi.org/10.1007/s11145-013-9495-7

Graham, S., Collins, A. A., & Rigby-Wills, H. (2017). Writing characteristics of students with learning disabilities and typically achieving peers: A meta-analysis. *Exceptional Children, 83*, 199-218. https://doi.org/10.1177/0014402916664070

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*, 445-476. https://doi.org/10.1037/0022-0663.99.3.445

Greenglass, E. R., Burke, R. J., & Konarski, R. (1997). The impact of social support on the development of burnout in teachers: Examination of model. *Work and Stress, 11*, 267-278.

Grünke, M., Boon, R. T., & Burke, M. (2015). Use of the randomization test in single-case research. *International Journal for Research in Learning Disabilities, 2*, 44-64.

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. *International Journal of Special Education, 30*, 137-150.

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.

Grünke, M., Wilbert, J., Tsiriotakis, I. K., & Agirregoikoa, A. L. (2017). Improving the length and quality of texts written by fourth graders with learning disabilities through a peer-tutoring graphic organizing strategy. *Insights into Learning Disabilities, 14*, 167-188.

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

Hartmann, D. P., Barrios, B. A., & Wood, D. D. (2004). Principles of behavioral observation. In S. N. Haynes & E. M. Heiby (Eds.), *Behavioral assessment* (pp. 108-127). Hoboken, NJ: John Wiley & Sons.

Hedges, L.V., Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods, 4*, 324-341. https://doi.org/10.1002/jrsm.1086

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.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in in special education. *Exceptional Children, 71*, 165-179. https://doi.org/10.1177/001440290507100203

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

Johnson, E. S., & Semmelroth, C. L. (2014). Special education teacher evaluation: Why it matters and what makes it challenging. *Assessment for Effective Intervention, 39*, 71-82.

Kinder, R. A. (2014). *1,000 awesome writing prompts*. North Charleston, SC: Create Space.

Kiuhara, S. A., Graham, S., & Hawken, L. S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology, 101*, 136-160. https://doi.org/10.1037/a0013097

Koutsoftas, A. D. (2016). Writing process products in intermediate-grade children with and without language-based learning disabilities. *Journal of Speech, Language & Hearing Research, 59*, 1471-1483. https://doi.org/10.1044/2016_JSLHR-L-15-0133

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. *What Works Clearinghouse*. http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kurihara, N. (2017). Do peer reviews help improve student writing abilities in an EFL high school classroom? *TESOL Journal, 8*, 450-470.

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.

Lobo, M. A., Moeyaert, M., Baraldi, C. A., & Babik, I. (2017). Single-case design, analysis, and quality assessment for intervention research. *Journal of Neurologic Physical Therapy, 41*, 187-197. https://doi.org/10.1097/NPT.0000000000000187

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.

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing. *The Journal of Special Education, 41*, 68-84. https://doi.org/10.1177/00224669070410020301

Moeyaert, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2014). From a single-level analysis to a multilevel analysis of single-case experimental designs. *Journal of School Psychology, 52*, 191-211. https://doi.org/10.1016/j.jsp.2013.11.003

Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2014). The influence of the design matrix on treatment effect estimates in the quantita-

tive analyses of single-subject experimental design research. *Behavior Modification, 38*, 665-704. https://doi.org/10.1177/0145445514535243

National Association of Colleges and Employers. (2012, November). *Job outlook 2011*. Retrieved from http://career.pages.tcnj.edu/files/2011/07/Job_Outlook_2011_Full_Report_PDF1.pdf

National Center for Education Statistics. (2012). *The nation's report card: Writing 2011* (NCES 2012-470). Washington, DC: U.S. Department of Education, Institute of Education Sciences.

National Commission on Writing. (2004). *Writing: A ticket to work … or a ticket out.* Retrieved from http://www.collegeboard.com/prod_downloads/writingcom/writing-ticket-to-work.pdf.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2010). Combining non-overlap and trend for single case research: Tau-U. *Behavior Therapy, 42*, 284-299. https://doi.org/10.1016/j.beth.2010.08.006

Puranik, C. S., Patchan, M. M., Lemons, C. J., & Al Otaiba, S. (2017). Using peer assisted strategies to teach early writing: Results of a pilot study to examine feasibility and promise. *Reading and Writing: An Interdisciplinary Journal, 30*, 25-50. https://doi.org/10.1007/s11145-016-9661-9

Rensing, J., Vierbuchen, M.-C., Hillenbrand, C., & Grünke, M. (2016). Implementing peer-assisted writing support in German secondary schools. *Insights into Learning Disabilities, 13*, 151-164.

Rietdijk, S., van Weijen, D., Janssen, T., van den Bergh, H., & Rijlaarsdam, G. (2018). Teaching writing in primary education: Classroom practice, time, teachers' beliefs and skills. *Journal of Educational Psychology, 110*, 640-663. http://dx.doi.org/10.1037/edu0000237

Saddler, B., & Asaro, K. (2007). Increasing story quality through planning and revising: Effects on young writers with learning disabilities. *Learning Disability Quarterly, 30*, 223-234. https://doi.org/10.2307/25474635

Saddler, B., & Asaro, K. (2008). The effects of peer assisted sentence combining practice on four writers with learning disabilities. *Learning Disabilities: A Contemporary Journal, 6*, 17-31.

Saddler, B., & Graham, S. (2005). The effects of peer-assisted sentence-combining instruction on the writing performance of more and less skilled young writers. *Journal of Educational Psychology, 97*, 43-54. https://doi.org/10.1037/0022-0663.97.1.43

Santangelo, T. (2014). Why is writing so difficult for students with learning disabilities? A narrative review to inform the design of effective instruction. *Learning Disabilities: A Contemporary Journal, 12*, 5-20.

Scardamalia, M., & Bereiter, C. (1986). Written composition. In M. Wittrock (Ed.), *Handbook on research on teaching* (pp. 778-803). New York, NY: MacMillan.

Scruggs, T., Mastropieri, M., & Casto, G. (1987). The quantitative synthesis of single subject research. *Remedial and Special Education, 8*, 24-33. https://doi.org/10.1177/074193258700800206

Scruggs, T. E., & Mastropieri, M. A. (1998). Synthesizing single subject studies: Issues and applications. *Behavior Modification, 22*, 221-242. https://doi.org/10.1177/01454455980223001

Skarr, A., Zielinski, K., Ruwe, K., Sharp, H., Williams, R. L., & McLaughlin, T. F. (2014). The effects of direct instruction flashcard and math racetrack procedures on mastery of basic multiplication facts by three elementary school students. *Education and Treatment of Children, 37*, 77-93.

Snodgrass, M. R., Chung, M. Y., Meadan, H., & Halle, J. W. (2018). Social validity in single-case research: A systematic literature review of prevalence and application. *Research in Developmental Disabilities, 74*, 160-173. https://doi.org/10.1016/j.ridd.2018.01.007

Stone, M., Friedlander, M. L., & Moeyaert, M. (2018). Illustrating novel techniques for analyzing single-case experiments: Effects of pre-session mindfulness practice. *Journal of Counseling Psychology, 65*, 690-702. https://doi.org/10.1037/cou0000291

Sundeen, T. H. (2014). Essay development for secondary students with learning disabilities: Graphic organizers for visualizing organizational patterns. *Rural Special Education Quarterly, 33*, 29-36. https://doi.org/10.1177/875687051403300305

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., … Wilson, B. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016 statement. *Archives of Scientific Psychology, 4*, 1-9.

UNESCO. (2011). *International standard classification of education*. Retrieved from http://uis.unesco.org/

Van den Noortgate, W., & Onghena, P. (2003a). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly, 18*, 325-346. https://doi.org/10.1521/scpq.18.3.325.22577

Van den Noortgate, W., & Onghena, P. (2003b). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instruments, & Computers, 35*, 1-10.

Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence Based Communication Assessment and Intervention, 2*, 142-151.

Vannest, K. J., & Hagan-Burke, S. (2009). Teacher time use in special education. *Remedial and Special Education, 31*, 126-142. https://doi.org/10.1177/0741932508327459

Vannest, K. J., Parker, R. I., & Gonen, O. (2011). Single case research: Web based calculators for SCR analysis (Version 1.0) [Web-based application]. College Station, TX: Texas A & M University. Available from singlecaseresearch.org.

Yarrow, F., & Topping, K. J. (2001). Collaborative writing: The effects of metacognitive prompting and structured peer interaction. *British Journal of Educational Psychology, 71*, 261-282. https://doi.org/10.1348/000709901158514