

To What Extent Do Certain Characteristics of a Child's Written Story Influence the Way It Is Rated? Insights Into Features Necessary for Supporting Struggling Writers

Matthias Grünke
Özlem Büyüknarci
University of Cologne

Jürgen Wilbert
University of Potsdam

Esther Breuer
University of Cologne

Many students exhibit a special type of learning disability in writing. Yet, teachers spend relatively little time helping children and youth with severe writing difficulties overcome their struggles or prevent emerging problems from becoming more severe. A major reason for this lack of attention is that many educators feel overburdened by the complex task of teaching students how to compose meaningful texts. The purpose of this explorative study was to determine the extent to which certain theory-based text- and person-related variables influence the way a written story is rated by experts and subsequently enable practitioners to make more informed decisions about where to start in supporting struggling writers. Sixty German secondary students (ages 10-13 years old) wrote stories that were evaluated by eight independent raters. Structural equation models were then used to examine the relationships between seven dependent variables and the quality of the texts. The analysis showed that rather short and illegible stories were generally rated unfavorable. Other factors (e.g., performance on alphabetic and copying tasks or spelling skills) also played a role, but to a lesser extent. These findings provide teachers with useful information about where to start when trying to prevent learning disabilities in writing. In particular, educators should focus on instructing students (a) how to brainstorm story ideas in order to enable them to produce texts of an acceptable length and (b) on how to improve the legibility of their handwriting.

Keywords: essay-writing skills, composition writing, appraisal of essays, handwriting, spelling, learning disabilities in writing

Insights on 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.

INTRODUCTION

Being able to put one's thoughts on paper is a critical skill for success in school and later professional endeavors. As soon as students have acquired basic writing skills, they are under constant demand to demonstrate their academic competence through the written texts they produce. In other words, possessing content knowledge is not enough. One needs also to be able to express the information in a form that is comprehensible to the reader.

Children learn these processes by composing simple stories at the end of their elementary education. Around the age of 10, they gradually develop the skills necessary to produce more refined texts, like compare/contrast essays, argumentative treatises, and other kinds of more complex genres (Becker-Mrotzek & Böttcher, 2012). Acquiring solid writing competencies is not only important in language classes, it is also vital for demonstrating knowledge on written exams, for example, in natural sciences. Without ample abilities to "think on paper," students are bound to fail in a great variety of subjects (Grünke & Leonard-Zabel, 2015).

Competence in composing text, or rather, the ability to transfer ideas or information into a linguistic form while following conventional patterns of achieving a communicative goal with a specific audience is as important as it is challenging (Graham & Harris, 2005). Indeed, according to Kame'enui and Simmons (1990), it is "the most complex of language skills..." [and] the last to develop in the sequence of listening, speaking, reading, and writing" (p. 420). Thus, text composition requires brain-based components such as intact attention and concentration, spatial and sequential production, memory, higher-order cognition, language competencies (including adequate vocabulary, grammatical structures, and orthography), as well as executive functioning (Feifer & De Fina, 2002).

In their well-known theory, Hayes and Flower (1980) describe the processes that a person undergoes in the course of writing, which can be roughly subdivided into idea generation, planning, translating, executing and reviewing. In later elaborations of this model, Hayes (1996, 2006, 2012) incorporated supplemental social and affective elements (e.g., problem solving, text interpretation, embedded reflection), further illustrating the intricacy that is involved in text production. Hayes especially stressed the importance of effectively coping with complexity in ever-changing new writing challenges.

Bereiter and Scardamalia (1987) provided an explanation of how novices can gradually master these highly demanding text production tasks, including the significance of increasing the fluency of the linguistic processes in order to lessen the demands on a person's memory system when it tries to hold and manipulate a large amount of information simultaneously. Along these lines, Berninger and Swanson (1994) proposed a model that highlights the mean-

ingful role that transcription skills play on the way to becoming a proficient writer, especially spelling and motor skills. Without fluency, the more mechanical aspects of execution make the writing process even more demanding, usually resulting in brief texts (Bereiter & Scardamalia, 1987; Kreiner, 1996; Van der Hoeven, 1999).

Several studies have suggested that students who do not demonstrate the above skills to a sufficient degree often produce texts that are considered inadequate by teachers and other expert raters. That is, if the ability to simultaneously hold in memory and manipulate a large amount of information and to cope with complexity in novel situations is low, the resulting texts are viewed as being of relatively meager quality (e.g., Berninger et al., 2010). When students make a comparatively large number of spelling errors, their writing products do not get rated very favorably either, (e.g., MacArthur & Graham, 1987). The same is true for students whose handwriting is poor and, therefore, submit texts that are difficult to decipher (e.g., Greifeneder, Zelt, Seele, Bottenberg, & Alt, 2012; Rose, 2009). Finally, if students lack the stamina or the knowledge to elaborate on a subject and hence compose proportionally short writing products, their texts are generally also evaluated rather negatively (Englert & Raphael, 1988).

A great number of children and youth have severe problems with mastering the task of composing a meaningful text. If these difficulties persist over time and reach a critical stage, they are seen as an expression of a specific learning disability in writing. Referring to the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013), Hahn and Morgan (2014) described these students as having writing skills (measured by individually administered tests) that fall substantially below those expected given the individual's chronological age, intelligence, and age-appropriate education (i.e., delays of two years or more). School-based epidemiological studies have documented that the prevalence of children and youth with learning disabilities in writing is even greater than those with dyslexia or dyscalculia (Mayes & Calhoun, 2006). Different studies suggest that up to 15% of all students exhibit this kind of disorder (Katusic, Colligan, Weaver, & Barbaresi, 2009).

Given these statistics, it is important that teachers support children and youth in improving their ability to first author simple stories and later attend to more sophisticated linguistic text forms. This is anything but easy, however. As stated above, writing is a very complex neurodevelopmental process, and its intricacy and unique nature make it difficult to determine exactly what to teach (Anderson & Keel, 2002; Kaufer, Hayes, & Flower, 1986; Ortega, 2009; van Wijk & Sanders, 1999). To educators, the task of imparting the knowledge and skills to compose a meaningful text, therefore, often appears overwhelming (Troia & Graham, 2003). As a consequence, this vital part of schooling gets pushed "to the dusty corners of the classroom" (Schlagal, 2013, p. 257).

To remedy this dilemma, we need reliable information about which problems and skills practitioners have to focus on in order to help students produce texts of acceptable quality.

PURPOSE OF THIS STUDY AND RESEARCH QUESTIONS

The theoretical models described above provide answers to the question of what skills and competencies are needed to effectively put one's thoughts on paper. However, we still do not know much about the extent to which a specific factor contributes to being able to compose an acceptable writing product. To our knowledge, the variables highlighted earlier have never been incorporated into one study aimed at providing information about the significance that each of these elements plays in creating a text that gets rated positively by experts.

From the major relevant theories of writing development and the corresponding empirical studies (see above), we have gained an understanding of the importance of sufficient skills in spelling, handwriting, producing longer pieces of work, and coping with complexity in novel situations for creating an appealing text. However, we do not know much about how these factors relate to each other and which factor(s) is/are considered more crucial than others.

To fill this void in the literature, this study was designed to determine the extent to which the aforementioned variables are relevant for producing an acceptable piece of writing and in which way(s) they are connected with each other. Story writing was chosen, because this is the genre that students use from the very beginning of their writing careers (see above). If they have not developed ample competencies in authoring a simple narrative, they will not be able to proceed to more advanced levels of writing (Graham & Harris, 2005).

In order to systematize the relevant variables, we divided them into two categories:

- (1) text-related factors that can be directly detected on the basis of a story produced by a student (by assessing the handwriting, by counting the number of words, and by calculating the ratio of orthographical errors) and
- (2) person-related factors (fluency in handwriting and ability to cope with complexity in novel situations) that have to be measured independently.

The study was explorative in nature and was designed to gain insight into the roles that these variables play for children aged 10 and a little older when they try to produce stories of acceptable quality (as rated by experts).

METHOD

Participants

Comprehensive school students. We assessed a total of 60 students – 34 boys and 26 girls. Most of them were 11 ($N = 29$) or 12 ($N = 23$) years

old. The remaining eight students were 10 ($N = 4$) and 13 ($N = 4$) years old. All attended an inclusive comprehensive school in a major city in Northrhine-Westfalia (Germany) that served grades 5 through 13. They were either in fifth ($N = 36$) or sixth grade ($N = 24$). According to their teachers, about 20% of the students came from an immigrant background. A little more than 5% of them had some type of special educational need (learning disabilities, behavioral problems, mild mental retardation, and/or autism spectrum disorders). Due to the lack of a suitable standardized test in German, we were not able to determine the number of girls and boys who exhibited a learning disability in writing or were at risk of developing one. However, the respective class teachers estimated that about a quarter of their students demonstrated major problems in composing text.

Raters. Each story that was subject to our analysis was evaluated by eight independent expert raters. For this purpose, we recruited 192 college students from a large German university (31 males and 161 females). All of them were enrolled in an undergraduate program in special education and had at least rudimentary experience in assessing and teaching children after completing an internship in a special school for slow learners for six weeks. Their ages ranged from 18 to 44 ($M = 23.38$; $SD = 3.88$).

Measures

Independent variable. We used a subtest from the General German Language Test (GGLT; Steinert, 2011), in which subjects were to write a story about a drawing that was presented to them. The drawing showed a man climbing up a ladder to a balcony, while a woman watched him from a window in a neighboring house. There were no time limits for finishing the task.

To evaluate the written products, we applied a German version of the National Association of Educational Progress (NAEP) Writing Scale for narrative texts (Canz, 2015). This evaluation grid developed by the NAEP (National Center of Education Statistics, 2011) is one of the most widely used and valid instruments for large-scale writing assessment (Lee & Stankov, 2012). The rubric consists of a set of six descriptions that characterize stories of different quality levels. They address common criteria for evaluating texts like coherence, word choice, and structure. When using the NAEP Writing Scale, a rater assigns one of the six quality levels to a given writing product (with "1" being the highest and "6" being the lowest ranking). To determine the extent of agreement between the eight judges, we applied Krippendorff's alpha (Krippendorff, 2011) as an inter-rater coefficient. The overall reliability was .78. Such a value can be considered as high (Krippendorff, 2004).

Dependent variables. The length of each story was evaluated by counting the number of words. In addition, the percentage of spelling errors was determined. To measure handwriting fluency, we asked the students to record the

letters of the alphabet as fast as possible (alphabet task), as well as to transcribe a given text as quick as possible (copying task). In each case, participants were stopped after 3 minutes. This twofold way to quantify handwriting fluency is rather common (Barnett, Henderson, Scheib, & Schulz, 2009). The reason behind this distinction is the fact that alphabet tasks are generally more strenuous than copying tasks, because the activation of long-term memory may increase the cognitive load of the child's working memory (Graham, Struck, Santoro, & Berninger, 2006; McCutchen, 2000). By considering both kinds of tasks, we tried to cover all aspects of handwriting fluency. Two special education students from a large German university independently counted the number of words, the spelling mistakes, the letters in the alphabet task, and the words in the copying task. The level of agreement between the two raters equaled 100%.

To evaluate the neatness of handwriting, we applied a rating protocol developed by Mahrhofer (2004) consisting of 13 subscales that provide information about different aspects of the neatness of a child's script, while taking the respective grade level into account. The results are then expressed through two indices – one for legibility and one for uniformity. According to Mahrhofer, the overall reliability of the instrument is .87 (Cronbach's alpha). The handwriting evaluation was performed independently by the first and fourth authors. Intra-class correlation coefficients (ICC) were .98 for legibility and .95 for uniformity.

Students' ability to cope with complexity in novel situations was measured using the German Number Combination Test (NCT; Oswald & Roth, 1987). The NCT is a trail-making instrument. On four different sheets, participants must connect randomly positioned numbers from 1 to 90 in the correct order as fast as possible within a 1-minute time limit. Results are expressed as the mean quantity of correctly connected numbers. According to the manual, the test-retest reliability of the NCT varies between .84 and .97. It correlates highly with standard psychometric tests of intelligence (Rammsayer & Stahl, 2007).

Procedures

Five examiners participated in administering the test battery in the school. One held a doctorate, one held a master's degree in special education, a third was a graduate student, and the remaining two were undergraduate students in special education at a large German university. The data was collected during the course of a school day during regular classroom activities.

The 192 undergraduate students described above graded the stories that the students had written about the drawing from the GGLT. These raters had previously been instructed in groups of about 20 each on how to use the NAEP Writing Scale during 45-minute training sessions. As part of the training, they were provided with several examples of very well, mediocre, and poorly

written narratives that had been appraised by experts. Subsequently, the raters were given five sample stories to evaluate for themselves. They received feedback on how well they took the assessment criteria of the NAEP Writing Scale into consideration. Each text was then independently rated by eight undergraduate students. The means of these appraisals served as a measure of narrative quality.

RESULTS

Table 1 provides descriptive values for all variables. In terms of skewness (Skew) and kurtosis (CK), all of them (except for *words copied*) show a normal distribution. Words copied were strongly compressed to the right side of the distribution (a prolonged tail to the left) with a steep peak. After squaring, the variable kurtosis was 0.1 and skewness -0.1. Therefore, the squared values were kept for all further analyses.

Table 1. Descriptives for NAEP Values and Indicators on the Text, and Student Level

Measure	<i>M</i>	<i>SD</i>	Median	Min	Max	Skew	Kurtosis	<i>SE</i>
NAEP value	2.83	1.06	2.75	1.00	5.57	0.33	-0.34	0.14
Total words	100.37	45.16	99.00	26.00	227.00	0.54	-0.19	5.83
Letters of the alphabet copied	48.11	16.47	48.33	0.00	76.00	-0.55	-0.01	2.13
Words copied	38.92	9.32	40.00	0.00	56.00	-1.30	3.45	1.20
NCT	29.63	5.83	28.50	16.50	45.00	0.31	0.03	0.75
Legibility	3.42	1.10	3.00	0.00	5.00	-0.47	0.18	0.14
Uniformity	3.08	0.76	3.00	2.00	5.00	0.40	-0.25	0.10
Error ratio	12%	9%	9%	0%	35%	0.81	-0.52	1%

The inter-correlations of all variables are shown in Table 2. All correlations to the NAEP value were as expected: The total number of words and legibility of the texts were strongly correlated, the copying speed of the letters of the alphabet, the copying speed of words, the uniformity of the writing, and the ratio of orthographic errors were moderately correlated, whereas NCT values were only weakly correlated. All correlations were in the expected direction.

Besides the NAEP values, the other measures showed a complex pattern of correlations. The two copying speed measures (letters of the alphabet and words) were strongly correlated as were the two measures for handwriting (legibility and uniformity). NCT values showed weak to mild correlations with the

word and alphabet copying measures and only a weak correlation to the ratio of orthographical errors. Orthographical error ratio was moderately to strongly correlated to legibility and uniformity of handwriting. This complex pattern of correlations among all variables made it necessary to analyze an integrated statistical model based on multiple regression.

Table 2. Correlations Between NAEP Values and Indicators on the Text and Student Level

Measure	1	2	3	4	5	6	7	8
1. NAEP value	-	-.63*	-.41*	-.40*	-.25	-.52*	-.37	.43*
2. Total words	-.63*	-	.53*	.59*	.13	.44*	.27	-.05
3. Letters of the alphabet copied	-.41*	.53*	-	.61*	.32	.32	.14	-.15
4. Words copied	-.40*	.59*	.61*	-	.29	.21	.05	-.12
5. NCT	-.25*	.13	.32*	.29*	-	.00	-.08	.01
6. Legibility	-.52*	.44*	.32*	.21	.00	-	.57*	-.46*
7. Uniformity	-.37*	.27*	.14	.05	-.08	.57*	-	-.41*
8. Error ratio	.43*	-.05	-.15	-.12	.01	-.46*	-.41*	-

Note. * $p < .05$. Indicators of significance above the diagonal are corrected for multiple tests following the Holm correction (Holm, 1979).

We set up three structural equation models (SEM) for an exploratory analysis of the model structure. The first model assumed that all manifested variables construed a single latent factor explaining the NAEP results (1-factor model). The second model assumed two separate latent factors explaining the NAEP results (2-factor model), where the first factor comprised all person-related manifest variables (alphabet letter copying speed, word copying speed, NCT results) and the second factor indicated all text characteristics (uniformity and legibility of the handwriting, ratio of spelling errors, and number of words). The third model assumed the two latent factors in a mediation context (2-factor mediation model), where the person factor influenced the text factor, which in turn influenced the NAEP results. That is, we assumed that there would be no direct connection between the person factor and the NAEP results.

The models were analyzed using the R package lavaan (Rosseel, 2012). We estimated the SEMs with a maximum likelihood procedure. Due to a high variation in variance, all exploratory variables were standardized for this analysis.

Figure 1. Three explorative structure equation models (SEM). Path values are standardized.

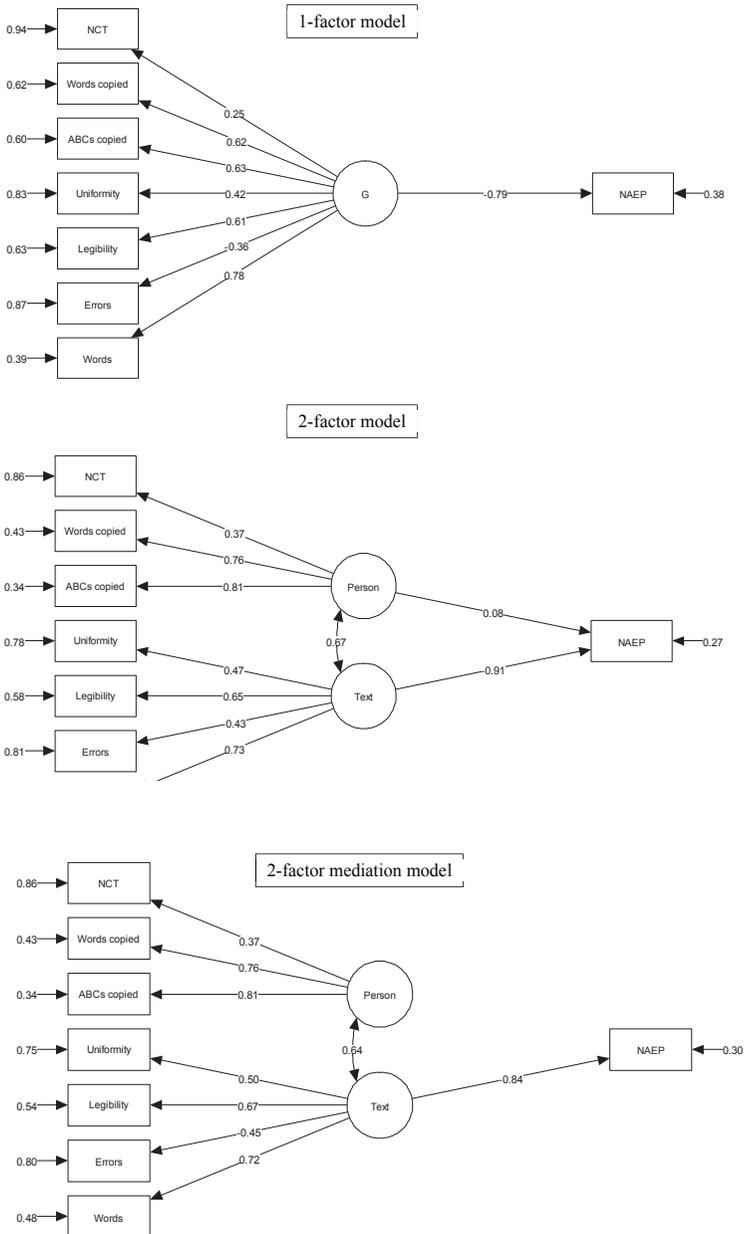


Table 3. Fit Indices of the Three Models

Model	χ^2	df	p	CFI	TLI	AIC	BIC	RMSEA	SRMR
1 factor	69.4	20	<.01	0.68	0.55	1272	1306	0.20	0.13
2 factors	52.9	18	<.01	0.77	0.65	1260	1298	0.18	0.11
2 factors mediation	53.0	19	<.01	0.78	0.68	1258	1294	0.17	0.11

All three models showed acceptable fit indices (see Table 3). However, the two 2-factor models produced the best model fit (1-factor vs. 2-factor model: $\Delta\chi^2(2) = 16.4, p < .01$). Moreover, the mediation model had a nearly identical fit to the model without mediation (2-factor mediation vs. 2-factor: $\Delta\chi^2(1) = 0.07, p > .78$). As the mediation model is sparser, it might be preferred over the model without mediation. This argument is strengthened by the fact that within the model without mediation, the latent person factor had virtually no direct influence on the NAEP results (see Figure 1).

DISCUSSION

Main Findings

In this study, we addressed the question of which text- and person-related factors seem to have a bearing on the quality of a story written by children of 10 years of age or a little older. The results of this explorative assessment indicate that length and legibility play prominent roles: The longer and the more legible a text is, the higher its (rated) quality. To a lesser extent, a low ratio of spelling errors and a high uniformity of handwriting are also associated with a higher level of excellence of a narrative. In addition, a child’s ability to rapidly copy the letters of the alphabet and different words is moderately, but positively related to text quality. The competency to cope with complexity in novel situations seems to play a rather negligible role in this context, however.

We incorporated all independent variables into three structural equation models: one with a general factor, one with two factors (one that represents person- and one that represents text-related aspects), and one mediation model that assumes that the person-related variables influence the text-related ones. Even though the fit indices of all three alternatives were not high enough to make strong statements about the connections between all variables, the mediation model contributed the most to answering our research question. The length as well as the legibility of a text seemed to influence the latent text factor to the greatest extent.

Limitations

Research studies of the writing process pose challenges for scholars that many other school-related topics do not (Grünke & Leonard-Zabel, 2015). Put-

ting one's thoughts on paper involves coordinating multiple cognitive, linguistic, and physical operations along with considering genre-specific conventions and keeping the intended audience in mind (Troia & Graham, 2003). To identify variables within this intricate process that can be reliably measured and put into context with each other is very demanding.

By nature, our study faced many of the same problems as other research in this area. To start with, we attempted to capture the quality of narratives by using a rating instrument that unfortunately cannot do justice to everything that constitutes text quality. Even though our raters achieved a remarkably high level of agreement, the fact remains that the validity of our instrument stays debatable, since there is no objective external criterion for the quality of a story. In addition, we used undergraduate students as expert raters. It remains unknown whether more experienced teachers would have made similar appraisals. Another limitation pertains to the possibility that the results of our analyses might at least partially reflect assessment biases of university students when rating the quality of narratives. For example, we cannot determine to what extent handwriting influences text quality itself or (just) the appraisal of text quality.

We chose to incorporate various student- and text-related variables derived from relevant theories of writing development. Even though we proceeded as objectively as possible, our decisions about which factors to include might not be completely beyond reproach. Thus, we cannot deny the possibility that other researchers might come to slightly different conclusions. Thus, we need to present our findings with a due degree of humility.

Finally, our results cannot be generalized to populations or text genres other than the ones targeted in the current study.

Practical Implications and Future Research

Even though our findings are somewhat limited – especially with regard to their validity and their generalizability – they can provide practitioners with valuable clues about where to start as they try to support struggling writers at the beginning of secondary school. When composing argumentative, informative, or explanatory texts on an advanced level, relatively short essays are often viewed as being of better quality than lengthier ones (Koutsoftas, 2014). However, this is not the case with narratives produced by children between 11 and 12 years old. As students develop initial composition skills while attending to stories as the most basic genre, length matters. Before they can revise and edit a writing product, students need to have enough material to work with. A third of all texts that our students handed in consisted of fewer than 80 words. According to a study by Rodríguez, Grünke, González-Castro, and Cerezo (2014), this is an alarmingly low value for children of that age. Struggling students need support in composing longer stories. In analyzing the process of writing, Rodríguez, Grünke, González-Castro, García, and Álvarez-García (2015) discovered that

children who come up with rather short texts invest remarkably little time in planning. They think about what to write while writing and revise their products before submitting them, but they usually do not sit back first to collect their thoughts and take notes about what to put on paper and about what they want to communicate to their readers. Having to mind their spelling and handwriting is challenging enough. If, in addition, students only start to contemplate the content of their story after they have started writing it, the assignment of producing a decent narrative will overburden them.

One way of helping children to elaborate more on their stories involves instructing them on how to execute effective planning strategies. Prewriting activities (like using graphic organizers) that serve the purpose of generating ideas before composing the actual story are very beneficial. In a study by Hennes, Büyüknarci, Rietz, and Grünke (2015), such an approach led to an effect size of Cohen's *d* of 0.41. According to a meta-analysis by Gillespie and Graham (2014) as well as one by Rogers and Graham (2008), prewriting activities not only increase the number of total words written, but can also positively impact the quality of an essay.

Graham (2010) highlighted the importance of effective handwriting instruction for enhancing the length and the quality of a story. An increased level of fluency enables children to invest a greater amount of mental effort into producing longer essays. If students learn not only to write faster but also to write more legibly (and more uniformly), this automation will further reduce the problem of having to attend to too many tasks at the same time. That is, a child who has invested a lot of time and energy into writing fluently and neatly can better attend to the content of a text. In a literature review, Graham (1999) identified several ways to improve a student's handwriting both effectively and efficiently. Even though the ratio of orthographical errors in a text was not associated with its quality to the same extent as length and legibility, spelling instruction would contribute to helping children focus on the content of their essay before, during, and after composing it. According to a literature review by Sayeski (2011), an explicit rule-based approach that provides extensive opportunities of practice with corrective feedback is generally the best option.

Future research should focus on replicating and expanding the results of this study through the use of different instruments and research designs. As mentioned, the NAEP Writing Scale for narrative texts is just one of several methods to capture the quality of a story. Prospective studies should consider different ways to measure how well a story is composed. In doing so, experts other than undergraduate students need to be used as raters as part of the variance in the current study may be due to differences in the raters.

REFERENCES

- American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. Washington, DC: Author.
- Anderson, D. M., & Keel, M. C. (2002). Using reasoning and writing to teach writing skills to students with learning disabilities and behavioral disorders. *Journal of Direct Instruction*, 2(1), 49-55.
- Barnett, A. L., Henderson, S. E., Scheib, B., & Schulz, J. (2009). Development and standardization of a new handwriting speed test. *British Journal of Educational Psychology*, 79(1), 137-157.
- Becker-Mrotzek, M., & Böttcher, I. (2012). *Schreibkompetenz entwickeln und beurteilen* [Developing and assessing writing skills]. Berlin, Germany: Cornelsen.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
- Berninger, V. W., Abbott, R. D., Swanson, H. L., Lovitt, D., Trivedi, P. Lin, S.-J., Gould, L., Youngstrom, M., Shimada, S., & Amtmann, D. (2010). Relationship of word- and sentence-level working memory to reading and writing in second, fourth, and sixth grade. *Language, Speech, and Hearing Services in Schools*, 41(2), 179-193.
- Berninger, V. W., & Swanson, H. L. (1994). Modifying Hayes' and Flower's model of skilled writing. In E. Butterfield (Ed.), *Children's writing: Toward a process theory of development of skilled writing* (pp. 57-81). Greenwich, CT: JAI Press.
- Canz, T. (2015). *Validitätsaspekte bei der Messung von Schreibkompetenzen* [Validity issues in the assessment of writing skills] (Doctoral dissertation). Retrieved from <http://edoc.hu-berlin.de/dissertationen/canz-thomas-2015-10-19/PDF/canz.pdf>.
- Englert, C. S., & Raphael, T. E. (1988). Constructing well-informed prose: Process, structure and metacognitive knowledge. *Exceptional Children*, 54(6), 513-520.
- Feifer, S. G., & De Fina, P. A. (2002). *The neuropsychology of written language disorders*. Middle-town, MD: School Neuropsych Press.
- Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children*, 80(4), 454-473.
- Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities: A review. *Learning Disability Quarterly*, 22(2), 78-98.
- Graham, S. (2010). Want to improve children's writing? Don't neglect their handwriting. *Education Digest*, 76(1), 49-55.
- Graham, S., & Harris, K. R. (2005). *Writing better: Effective strategies for teaching students with learning difficulties*. Baltimore, MD: Brookes.
- Graham, S., Struck, M., Santoro, J., & Berninger, V. (2006). Dimensions of good and poor handwriting legibility in first and second graders: Motor programs, visual-spatial arrangement, and letter formation parameter setting. *Developmental Neuropsychology*, 29(1), 43-60.
- Greifeneder, R., Zelt, S., Seele, T., Bottenberg, K., & Alt, A. (2012). Towards a better understanding of the legibility bias in performance assessments: The case of gender-based inferences. *British Journal of Educational Psychology*, 82(3), 361-374.
- Grünke, M., & Leonard-Zabel, A. M. (2015). How to support struggling writers: What the research stipulates. *International Journal of Special Education*, 30(3), 30(3), 137-150.
- Hahn, L. G., & Morgan, J. E. (2014). ADHD and learning disorders. In L. A. Adler, T. J. Spencer, & T. E. Wilens (Eds.), *Attention-deficit hyperactivity disorder in adults and children* (pp. 123-138). Cambridge, UK: Cambridge University Press.
- Hayes, J. R. (1996). A new model of cognition and affect in writing. In M. Levy & S. Ransdell (Eds.), *The science of writing* (pp. 1-27). Hillsdale, NJ: Erlbaum.

- Hayes, J. R. (2006). New directions in writing theory. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 28-40). New York, NY: Guilford.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369-388.
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Erlbaum.
- 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 to use the story maps strategy. *Insights on Learning Disabilities, 12*(1), 35-56.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70.
- Kame'enui, E. J. & Simmons, D. C. (1990). *Designing instructional strategies: The prevention of academic learning problems*. Columbus, OH: Merrill.
- Katusic, S. K., Colligan, R. C., Weaver, A. L., & Barbaresi, W. J. (2009). The forgotten learning disability: Epidemiology of written-language disorder in a population-based birth cohort (1976-1982). *Pediatrics, 123*(5), 1306-1313.
- Kaufer, D. S., Hayes, J. R., & Flower, L. S. (1986). Composing written sentences. *Research in the Teaching of English, 20*(2), 121-140.
- Koutsoftas, A. (2014, August). *Relationships among language measures and writing quality in intermediate grade children*. Paper presented at the Writing Research Conference of the European Association of Learning and Instruction (EARLI) in Amsterdam, The Netherlands.
- Kreiner, D. S. (1996). Effects of word familiarity and phoneme-grapheme polygraphy on oral spelling time and accuracy. *Psychological Record, 46*(1), 49-70.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.
- Krippendorff, K. (2011). *Computing Krippendorff's alpha-reliability*. Retrieved from http://repository.upenn.edu/asc_papers/43
- Lee, J., & Stankov, L. (2012). Special issue on non-cognitive psychological processes and academic achievement. *Educational Psychology, 32*(4), 551-552.
- MacArthur, C. A., & Graham, S. (1987). Learning disabled students' composing under three methods of text production: Handwriting, word processing, and dictation. *The Journal of Special Education, 21*(3), 22-42.
- Mahrhofer, C. (2004). *Schreibenlernen mit graphomotorisch vereinfachten Schreibvorgaben* [Learning to write under grapho-motorically simplified writing requirements]. Bad Heilbrunn, Germany: Klinkhardt.
- Mayes, S. D., & Calhoun, S. L. (2006). Frequency of reading, math, and writing disabilities in children with clinical disorders. *Learning and Individual Differences, 16*(2), 145-157.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist, 35*(1), 13-23.
- National Center of Education Statistics. (2011). *Writing 2011*. Washington, DC: Department of Education, Office of Educational Research and Improvement.
- Ortega, L. (2009). *Understanding second language acquisition*. London, UK: Hodder Education.
- Oswald, W. D., & Roth, E. (1987). *Zahlen-Verbindungs-Test* [Number Combination Test]. Göttingen, Germany: Hogrefe.
- Rammsayer, T. H., & Stahl, J. (2007). Identification of sensorimotor components accounting for individual variability in Zahlen-Verbindungs-Test (ZVT) performance. *Intelligence, 35*(6), 623-630.

- Rodríguez, C., Grünke, M., González-Castro, P., & Cerezo, R. (2014, August). *Writing composition, product and process measures in students with attention-deficit/hyperactivity disorder*. Paper presented at the Writing Research Conference of the European Association of Learning and Instruction (EARLI) in Amsterdam, The Netherlands.
- 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 nonlabeled peers in the ability to compose texts? *Learning Disabilities: A Contemporary Journal*, 13(2), 191-213.
- Rogers, L. A., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology*, 100(4), 879-906.
- Rose, J. (2009). *Identifying and teaching children and young people with dyslexia and literacy difficulties*. London, UK: DCSF Publications.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36.
- Sayeski, K. L. (2011). Effective spelling instruction for students with learning disabilities. *Intervention in School and Clinic*, 47(2), 75-81.
- Schlagal, B. (2013). Best practices in spelling and handwriting. In S. Graham, C. A. MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (pp. 257-283). New York, NY: Guilford.
- Steinert, J. (2011). *Allgemeiner Deutscher Sprachtest* [General German Language Test]. Göttingen, Germany: Hogrefe.
- Troia, G. A., & Graham, S. (2003). Effective writing instruction across the grades: What every educational consultant should know. *Journal of Educational and Psychological Consultation*, 14(1), 75-89.
- Van der Hoeven, J. (1999). Differences in writing performance: Generating as indicator. In M. Torrance & D. Galbraith (Eds.), *Studies in writing Vol. 4: Knowing what to write* (pp. 65-77). Amsterdam, The Netherlands: Amsterdam University Press.
- Van Wijk, C., & Sanders, T. (1999). Identifying writing strategies through text analysis. *Written Communication*, 16(1), 52-76.

AUTHORS' NOTE

Correspondence concerning this article should be addressed to Matthias Grünke, University of Cologne, Department of Special Education and Rehabilitation, Klosterstr. 79b, 50931 Cologne, Germany. Contact: matthias.gruenke@uni-koeln.de.