The Relevance of Executive Functions for Writing Competence in Skilled Writers

Julie Wolf, Rebecca Maria Kreutz, Ann-Kathrin Hennes & Alfred Schabmann

Department of Special Education and Rehabilitation, Faculty of Human Sciences, University of Cologne | Germany

Abstract: The relevance of the three core executive functions for text quality and individual text features had already been shown in a number of studies with novice writers. However, experimental data on the influence of executive functions in adult, skilled writers are rare. This cross-sectional study thus investigated how inhibition, updating, and shifting are related to text quality and relevant individual text features in skilled writers. Sixty-three adult, skilled writers wrote an explanatory text. The text product was used to analyze text quality (as perceived by the recipient) and individual text features (information content, cohesion, lexical diversity, appropriate words, syntactic complexity). Executive functions were assessed with standardized tests. Analyses highlighted the importance of conveying cohesive information when writing explanatory texts. Regarding the role of executive functions, inhibition significantly influenced information content and cohesion; updating influenced the use of appropriate words; and shifting influenced perceived text quality. These findings emphasize that executive functions play an important role in writing not only for novice writers but across the lifespan, whereby their influence changes with development and expertise.

Keywords: executive functions, text quality, text features, skilled writing

journal of WRITING RESEARCH

Wolf, J., Kreutz, R.M., Hennes, A.K., & Schabmann, A. (Accepted for publication in 2025). The relevance of executive functions for writing competence in skilled writers. Journal of Writing Research, *volume(issue),* ##-##. DOI: xx

Contact: Julie Wolf, Department of Special Education and Rehabilitation, Faculty of Human Sciences, University of Cologne, KlosterstraЯe 79b, 50931 Cologne | Germany – julie.wolf@uni-koeln.de

1. Introduction

Composing texts is a complex and cognitively demanding task that requires the integration of multiple cognitive processes and skills. Therefore, executive functions (EFs) play a central role in its accomplishment (Cordeiro et al., 2020).

1.1 Executive functions

EFs refer to a set of domain-general cognitive processes that enable goal-directed behavior, decision making, and problem solving (Diamond, 2013). The cognitive processes most frequently postulated when defining EFs are the three core low-level functions described by Miyake et al. (2000): Inhibition is the ability to consciously suppress dominant or automatic responses; updating refers to the dynamic manipulation of information in working memory; and shifting is the ability to switch attention between tasks, strategies, or mental sets. Inhibition, updating and shifting are particularly involved in complex, long-term cognitive activities that demand intensive monitoring and in which current and future goals must be kept up-to-date and various processes and skills must be managed and coordinated (Diamond, 2013). These requirements also apply when composing texts.

1.2 Writing

Composing high-quality texts is crucial for academic and professional success as well as for social participation (Crossley & McNamara, 2016). The underlying writing competence relies on low-level transcription skills (handwriting fluency and spelling) as well as higher-level text generation skills, including language skills (e.g., vocabulary knowledge) and more advanced cognitive skills (e.g., perspective taking) (Berninger & Winn, 2006; Hayes, 2012; Kim & Graham, 2022).

Assessing writing competence is a challenging task (Ruffini et al., 2024). In both, educational and research contexts, writing tasks of different genres are often assigned (Feenstra, 2021; Grabowski et al., 2014). The quality of the resulting texts, which integrates various skills required to different degrees depending on the genre, can then be evaluated using different approaches (Grabowski et al., 2014). One approach involves holistic judgements by human raters, who evaluate texts as a whole. This type of evaluation always reflects the perceived text quality (pTQ) from the recipient's perspective. Alternatively, text features that provide a more specific indication of the use of higher-level text generation skills can be analyzed at the macro- and micro-level of a text – either by human raters or automated tools (Dockrell & Connelly, 2021; Donovan & Smolkin, 2006). At the macro-level, these text features include the information content and the organization of ideas.

The *information content* of a text refers to the amount and relevance of the information (respectively ideas) that a text conveys. Composing high-quality texts involves generating and selecting this information purposefully, based on the

writing goal and the audience (Becker-Mrotzek et al., 2014; Bereiter & Scardamalia, 1987; Hayes, 2012; Hennes, 2020). Writers must select which information must be given explicitly and which (implicit) information readers can infer independently based on the anticipated prior knowledge of the intended audience (Hennes, 2020). For the genre of factual texts, which include among others argumentative, instructional and explanatory texts (Martin & Rothery, 1980), empirical studies have confirmed a close relationship between information content and pTQ (e.g. Crossley & McNamara, 2010, 2016; Grabowski et al., 2014).

The *organization of ideas* refers to the logical arrangement and connection of sentences and paragraphs to express ideas in a way the reader can easily understand and follow. When ideas and information are clearly connected both within and between paragraphs, a text is considered coherent (NAEP, 2011). Achieving coherence involves linking parts of a text through lexical or grammatical linguistic devices, known as cohesion, while also depending on the reader's prior knowledge and abilities (Struthers et al., 2013). Insufficient cohesion can make it challenging for readers to recognize the systematic relationship between common lexical items and thus impair comprehension (McNamara et al., 2010). Empirical studies have also shown that cohesion is closely related to pTQ in both narratives and factual texts (e.g. MacArthur et al., 2019; Philippek et al., 2025).

At the micro-level of a text, the deliberate use of *linguistic devices* helps to ensure that the intended message is clear and that the genre-specific goal of the text is supported (NAEP, 2011). Analyzing these linguistic devices enables conclusions to be drawn about language skills – especially in the area of vocabulary and sentence structure (Grabowski et al., 2014; Hennes, 2020; NAEP, 2011). In this context, examining lexical diversity, the use of appropriate words, and the complexity and variation of sentence structures—whose relevance may vary depending on the text genre—is well-established in writing research. While empirical studies have shown that lexical diversity is a reliable predictor of text quality especially in narratives (Olinghouse & Wilson, 2013), the use of appropriate words is above all related to pTQ in factual texts (Crossley & McNamara, 2010; Mathiebe, 2019). Syntactic complexity is relevant for pTQ of narratives as well as factual texts (Beers & Nagy, 2009; McNamara et al., 2010).

1.3 Executive functions and writing

Theoretical models of writing assume that EFs are crucial for both novice and skilled writers. According to the *Not-so-Simple View of Writing model* (Berninger & Winn, 2006), in the early stages of writing development, the limited resources of working memory and EFs are primarily required for transcription processes. In the course of development, these processes become automatized, and the freed-up cognitive resources can be used for higher-level text generation skills and control of the writing process. Consequently, skilled writers are increasingly able to manage the

processes of writing simultaneously and apply the communicative writing strategy of *knowledge transforming* (as opposed to *knowledge telling* used by novice writers), which involves composing texts purposefully for a specific readership and giving coherence to a text (Bereiter & Scardamalia, 1987). Similarly, in the *Direct and Indirect Effects Model of Writing* (Kim & Graham, 2022), EFs are considered foundational to the writing process, influencing both low-level transcription skills and higher-level text generation skills, and thereby indirectly affecting overall text quality.

These model theoretical assumptions have already been supported by several studies focusing on novice writers.

The majority of these studies showed that *inhibition* plays a significant role in writing. Altemeier et al. (2008) found that inhibition influenced students' written expression (including word fluency, sentence combining, and paragraph writing) from first through fifth grade. Rocha et al. (2022) found that inhibition in the fourth grade predicted pTQ in fifth grade. Moreover, inhibition affected transcription skills in students between the first and fifth grades (Altemeier et al., 2008; Drijbooms et al., 2015). No direct effects were observed on the information content and syntactic complexity in students aged 9–12 years. However, indirect effects were identified, showing that handwriting fluency influenced both information content and syntactic complexity in these students (Drijbooms et al., 2015; Zahra et al., 2023).

Studies on the role of *updating* in writing have consistently shown that updating contributes to the variance in transcription skills of students in grades 3 through 5 (Balioussis et al., 2012; Drijbooms et al., 2015). However, mixed results were obtained regarding pTQ: While some studies found that updating explains parts of the variance in pTQ in written compositions of students aged 6 to 7 years (Bourke & Adams, 2003) and 11 years (Connelly et al., 2012), Rocha et al. (2022) found that updating in the fourth grade did not predict pTQ in fifth grade. While Balioussis et al. (2012) found effects on syntactic complexity in third and fifth grade students, Bourke and Adams (2003) found effects on lexical diversity and coherence, but not on syntactic complexity in 6- to 7-year-old students. Other studies also found no direct effect of updating on syntactic complexity, although they did identify indirect effects through handwriting fluency. These indirect influences affected both syntactic complexity and information content in 9- to 12-year-old students (Drijbooms et al., 2015; Zahra et al., 2023).

The role of *shifting* in writing has been less studied, however, and the existing research revealed partially inconsistent results (Ruffini et al., 2024). Rocha et al. (2022) found that shifting in the fourth grade predicted pTQ in fifth grade, while Altemeier et al. (2008) observed a significant effect on written expression only in third graders when assessing students from grades 1 through 5. Moreover, Balioussis et al. (2012) found that shifting predicted handwriting fluency and syntactic complexity in third through fifth graders. Overall, the study results

indicate that the core EFs play a different but fundamental (not yet fully understood) role for novice writers. When transcription skills are not yet automatized, inhibition is needed to suppress incorrect letters, motor movements, or spelling patterns, and updating helps to keep a phonological form active until an orthographic rule is applied or a motor integration has been completed (Altemeier et al., 2008).

In skilled writers, when transcription skills are automatized, the freed-up cognitive resources can be used for higher-level text generation skills and are thus likely to still play a crucial role in the writing process. However, as experimental data about the influence of EFs in adult, skilled writers are rare, assumptions can only be deduced from indirect, related empirical findings (e.g., studies with novice writers) or theoretical descriptions of the cognitive demands of writing (Olive, 2021). EFs might be needed by these writers to select relevant information, structure texts in a coherent way, and coordinate the writing process (Kellogg, 2008; Olive, 2021). In particular, inhibition might be important, as it enables writers to suppress irrelevant information and focus on what is essential (Hooper et al., 2021; Olive, 2021). Updating might be needed to refine information or modify initial plans and ideas while composing texts (Ruffini et al., 2024). When applying the writing strategy *knowledge transforming,* it might also be needed to store and update the communicative goal and the reader's representation during the writing process (Kellogg, 2008). Shifting might allow skilled writers to flexibly switch between different ideas of a global text and manage the activities of the writing process simultaneously (Hooper et al., 2021; Olive, 2021).

1.4 Current study

Since previous studies had examined the relevance of EFs for writing in children (e.g., novice and developing writers), this cross-sectional study aimed to expand our understanding of the complex construct of writing competence to skilled writers. More specifically, with this study, we sought to analyze the role of EFs in skilled writers to better understand how the influence of these cognitive abilities changes over the course of development and with increasing writing expertise.

For the purpose of this study, we chose to focus on the EFs inhibition, updating, and shifting for three reasons. First, this three-part model of relatively circumscribed, lower-level EFs is used by many researchers in the field of cognitive psychology and provides a clear conceptual basis for defining and operationalizing the complex construct of EFs (Miyake et al., 2000). Second, these EFs are probably involved in more complex higher-level EFs like planning and monitoring, which are also important for writing (Hooper et al., 2021). Third, the role of these EFs for novice writing has already been shown in several studies.

In terms of writing competence, we measured both the pTQ and individual text features that provide a more specific indication of the use of higher-level text

generation skills. We did this by analyzing the text products of explanatory writing tasks from skilled writers (for details, see the Method section).

Based on these considerations, the following research questions emerged:

1. *Which text features are relevant for the perceived text quality of explanatory texts written by skilled writers?*

Based on the previous study results, we hypothesized that pTQ would be influenced by information content, cohesion, the use of appropriate words, and syntactic complexity. Moreover, we assumed that the text features relevant for pTQ would be interrelated and jointly contribute to the perception of text quality.

2. *How do the three core executive functions influence perceived text quality and the relevant text features of explanatory texts written by skilled writers?*

According to findings from studies with novice writers and theories on the cognitive demands of writing, we assumed that the three core EFs would also be required by skilled writers and would be deployed for higher-level text generation skills. For this reason, we expected that EFs would influence text features relevant for pTQ (see research question 1) and pTQ itself.

2. Method

2.1 Participants

This study included a total of 63 adult participants (students of humanities and administrative staff) recruited via email invitations from the University of Cologne. The proportion of positive responses was approximately 6%. The proportion of women in the sample was around 70%, which reflects the general gender distribution within the target populations (Kortendiek et al., 2022; p. 125)

All participants met the following inclusion criteria (see Table 1): They had to be at least 18 years old, native German speakers, have no diagnosed reading and/or spelling disorder, and possess at least a high school diploma. Moreover, the participants had to write longer texts for academic or professional purposes several times a week, regularly write on a keyboard, and their typing skills had to be automatized.

The automatization of typing skills was assessed by measuring typing fluency and typing error rates. Writers with automatized typing skills have a higher typing fluency and lower error rates than novice writers. In this study, participants were excluded if their typing fluency was below 26 words per minute or their error rate exceeded the following thresholds: 3.72% substitution errors, 1.29% deletion errors, or 0.80% insertion errors (cut-off values for slow typists from Dhakal et al., 2018).Participants provided informed consent before participating in this study; the

data were collected completely anonymously. To compensate for their time and effort, participants received some financial compensation.

Table 1. Characteristics of the participants

Variable	Descriptive statistics
Age	M = 35.86 (range: 19–62) years, SD = 14.98
Gender	Females: 44
	Males: 19
Profession	Students: 33
	Administrative staff of the university: 30
Highest educational qualification	High school diploma: 8
	University entrance qualification: 32
	University degree: 23
Frequency of writing longer texts	Daily: 49; several times a week: 14
Writing medium used	Exclusively keyboard: 10
	Keyboard > pen: 43
	Keyboard = pen: 10
Typing fluency	M = 61.29 (range: 34–116) words per minute, SD = 15.59
Words with substitution errors	M = 0.02% (maximum: 0.25%), SD = 0.04
Words with omission errors	M = 0.01% (maximum: 0.10%), SD = 0.02
Words with insertion errors	M = 0.02% (maximum: 0.21%), SD = 0.04

Note. M = mean; SD = standard deviation.

2.2 Measures

A self-report questionnaire was used to gather participants' demographic data and information on their writing behavior. Participants' typing skills were assessed with a sentence copying task and the EFs with standardized tests. Writing competence was assessed through an explanatory task (adapted from Schreibkompetenztest 4-9" in development, see https://idsl2.phil-fak.uni-koeln.de/institut/personen/lehrendenseiten/prof-dr-joerg-jost/forschungsprojekte-research-projects/schreibkompetenztest-4-9-writing-test-4th-9th-grade), which was designed to be culturally neutral, independent of prior knowledge, and cognitively stimulating (Jost, 2022).

Participants first watched a non-verbal video explaining the functioning of a bubble gun and were then given the following prompt: "Please write an explanatory text for an adult who has not seen the video, explaining how a bubble gun works." They were allowed to watch the video again as often as necessary while writing their text in case they needed to recall the steps. Explanatory texts (as a subtype of factual texts) were chosen due to their importance in education, workplaces, and society (Fang et al., 2020) and because they demand sophisticated knowledge transforming (Beauvais et al., 2011). Investigating the role of EFs is easier with such texts than with those from simpler genres (e.g., narratives), which can be approached with knowledge-telling strategies and therefore require less cognitive effort (Vieira et al.,

2023). Spelling errors in the texts were corrected before analyses in analogy to MacArthur et al. (2019).

Perceived text quality

To capture pTQ, the method of comparative judgment was conducted, which addresses issues of reliability and validity in text evaluation by employing direct holistic comparisons of texts by a group of independent raters. Texts are randomly paired off and raters must decide which of the two texts is better. Based on these decisions, a logit score per text is determined with a logistic model. This score indicates the probability of winning a comparison with a reference text and can be used to develop a ranking (Lesterhuis et al., 2016). The scale separation reliability (analogue to Cronbach's alpha) determines the reliability of the estimated logit scores (Jones & Karadeniz, 2016). Satisfactory values of reliability ($r > .7$) can be achieved with 10 to 14 comparisons per text. With 15 or more comparisons per text, good values for convergent validity can be attained both for experts and naïve raters (Verhavert et al., 2019).

Comparative judgments were conducted with the online tool Comproved (www.comproved.com). The texts were evaluated by naïve raters (third-semester students at the faculty of human sciences). Each rater received eight pairs of texts and was asked to decide which of the two texts better described how the bubble gun works.

The assessments were conducted in two sessions. In the first session, 30 texts (from the first data collection, see Procedure) were evaluated by 163 raters. A total of 887 comparisons were made; each text was rated at least 29 times. The scale separation reliability was $r = .83$. In the second session, 33 texts (from the second data collection) and two anchor texts from the first session (which were included in order to rank all texts on a common scale) were evaluated by 97 raters. A total of 565 comparisons were made; each text was rated at least 15 times. The scale separation reliability was $r = .81$. As the calculated logit scores of both evaluation sessions are only informative within the same ranking, the logit scores of the first session were converted into grade scores on a scale of 0–100 based on the idea of fair averages (Eckes, 2019). With the help of the grade scores of the anchor texts of the first session, the logit scores of the second session were also converted.

Text features

The text product was also used to measure various text features.

Information content. The first step to assess information content was to develop a catalogue of criteria that defined the information absolutely necessary for a complete understanding of how the bubble gun works (procedure adapted from Grabowski et al., 2014). This catalogue was developed by three linguistic experts who identified every possible piece of information and decided by consensus

whether it was necessary for a comprehensible explanation of the functioning of a bubble gun. This process resulted in a catalogue of criteria with 14 pieces of information (see appendix). Then, each text was analyzed by another two linguistic experts to determine whether it contained the information listed in the criteria catalogue. One point was awarded for each piece of correctly given information. The interrater reliability (IRR) of the individual criteria was between $\kappa = .81$ ($p < .001$) to $\kappa = 1$ ($p < .001$). In the few cases of disagreement, the case was discussed until a consensus was reached. The final score for information content was the sum of the correctly provided information.

Cohesion. Since cohesion can be measured at the surface of a text, it serves as a good indicator of coherence (Struthers et al., 2013) and was therefore used as a proxy for the organization of ideas in this study. Cohesion was assessed by two linguistic experts who used a rating system adopted from Philippek et al. (2025). The assessment focused on local grammatical and semantic cohesion errors that could impair comprehension, such as the absence of necessary references or the connection of unrelated elements. The IRR of cohesion errors was calculated ($\kappa = .79$; $p < .001$) and cases of disagreement were discussed until a consensus was reached. Then, cohesion errors were divided by the number of clauses in the text, which were also counted by the two raters with an IRR of $\kappa = .90$ ($p < .001$). Subsequently, the proportion of local cohesion was calculated (1 - proportion of cohesion errors).

Lexical diversity. Lexical diversity is usually assessed quantitatively, such as with the type-token ratio (TTR), which divides unique words (types) by total words (tokens) in a text. To minimize text length bias, this study used the measure of textual lexical diversity (MTLD), known to be less affected by text length (Koizumi & In'nami, 2012). The MTLD represents the average length of successive word strings in a text that retain a given TTR value. McCarthy and Jarvis (2010) showed that TTR curves tend to stabilize at around 0.72. The number of times the text reaches this TTR value is counted from the beginning to the end of the text. Then, the mean word count is calculated. This is done by counting the number of tokens and dividing by the number of times the text reaches the specified TTR value. Once this first cycle has been completed and an initial MTLD score has been calculated, the entire text is analyzed again in reverse order, resulting in another MTLD score. The final MTLD score is the average of the forward and reverse MTLD scores. This measure is highly correlated with other measures of lexical diversity and therefore has satisfactory convergent validity. Furthermore, MTLD is reliable in that shorter sections of a text have similar MTLD scores to the whole text, and the MTLD scores of these text sections do not correlate with text length (McCarthy & Jarvis, 2010).

Appropriate words. First, a list of appropriate synonyms for the required technical terms was compiled by two linguistic experts. Subsequently, all texts were assessed for inappropriate words (inappropriate synonyms and words in the text

that did not fit into the context) by another two linguistic experts (procedure adopted from Philippek et al., 2025). For this measurement, the IRR was $\kappa = 0.62$ ($p < .001$). Discrepancies were discussed until consensus was reached. The proportion of appropriate words was then calculated (1 - number of inappropriate words/ number of words).

Syntactic complexity. To assess syntactic complexity, the number of words per clause was calculated, which is one of the most commonly used measures (Beers & Nagy, 2009). To do this, the number of words in the text (without heading) was divided by the number of clauses.

Typing skills

A sentence copying task (adapted from Olinghouse, 2008) was used to assess participants' typing skills. Participants were visually presented with a sentence containing all the letters of the German alphabet ("Franz jagt im komplett verwahrlosten Taxi quer durch Bayern [Franz races across Bavaria in a completely derelict taxi]"). They were asked to read this sentence and then repeatedly type it as accurately and quickly as possible for 60 seconds. Typing fluency was determined by counting all copied words. Error rates were calculated by counting all words with substitutions, omissions, and insertions, using the Microsoft Office Word spellchecker and then dividing them by the number of all copied words.

Executive functions

The three core EFs were measured with standardized tests.

Inhibition. In this study, the "go/no-go task" from the German computer-based test battery for attention testing [Testbatterie zur Aufmerksamkeitsprüfung] (TAP; Zimmermann & Fimm, 2009) was used for testing inhibition. Here, squares with different filling patterns represent the "go" and "no-go" stimuli. Two out of five squares are previously defined as "go" stimuli. A sequence of squares with different filling patterns then appears in the center of a screen. The participants are to react as quickly as possible by clicking the mouse when the "go" stimulus appears and to show no response to "no-go" stimuli. A total of 60 stimuli (24 "go" stimuli) are presented in this specific task. Two relevant measures were extracted from this task: Inhibition speed was determined by the average reaction time to the "go" stimuli. Outliers were excluded when calculating this value with any reaction time that was more than 2.35 times the standard deviation above or below the individual's mean reaction time. The odd–even reliability of this measure is $r = .93$. Inhibition accuracy was assessed based on the number of correct responses with an odd–even reliability of $r = .67$ (Zimmermann & Fimm, 2009).

Updating. To assess updating, digit span tasks are often used (e.g. Bourke & Adams, 2003; Connelly et al., 2012). However, they tend to measure both passive storage and active updating processes of working memory (Cornoldi & Vecchi,

2003; St Clair-Thompson & Wen, 2021). For this reason, an N-back paradigm, more specifically the N-back task from the subtest "working memory" of the TAP (Zimmermann & Fimm, 2009), was used in this study, as it captures the isolated active process of updating more clearly (Owen et al., 2005; Ruffini et al., 2024). In this task, participants are presented with a sequence of single-digit numbers on a screen. Each number must be checked for a match with the penultimate number. If the numbers match, the participants are to react as quickly as possible by clicking the mouse. The test comprises a total of 100 stimuli (15 critical stimuli) presented at a rhythm of 3 seconds. Accuracy and speed were measured for updating in the same way as for measuring inhibition. The odd–even reliability is r = .85 for the speed and r = .74 for the accuracy measurement (Zimmermann & Fimm, 2009).

Shifting. A commonly used task to assess shifting is the Trail Making Test (TMT), consisting of two parts: In Part A, participants are required to connect a series of irregularly spaced numbers (1–25) on a sheet of paper with a line in ascending order. In Part B, participants are asked to connect numbers in ascending order and letters in the order of the alphabet, alternating between numbers and letters (e.g., 1-A-2-B). In this study, the TMT was administered in a standardized manner (Lezak, 2004), meaning that Part A preceded Part B, and that the correction of errors was included in the processing times. A relevant parameter for measuring shifting is the difference between the processing time in Part B versus Part A. The smaller the difference value is, the better the shifting ability. The reliability coefficients of this difference score vary considerably between studies, with most being r > .60 (Lezak, 2004).

Formation of the constructs of executive functions. Factor analyses (principal component analysis, varimax rotation) were conducted to reduce and summarize the EF data. Therefore, the data's suitability for factor analysis was checked. The Kaiser-Meyer-Olkin (KMO) test (KMO = .62) and Bartlett's test ($p < .001$) confirmed that the variables were appropriate for factor analysis. Using the criterion of eigenvalues greater than 1, the principal component analysis extracted two factors (see Table 2).

Table 2. Rotated factor loadings for all executive function variables

Variable	Component	
	1	2
Inhibition speed	.79	.26
Inhibition accuracy	.90	.00
Updating speed	.19	.81
Updating accuracy	.05	.86
Shifting	.32	.32

As the variable shifting did not show a strong loading on one of the two factors, it was considered as a single variable in the following analyses, and the factor analysis was rerun with the remaining four variables (inhibition speed, inhibition accuracy, updating speed, updating accuracy). These data were again categorized as suitable for a factor analysis based on the KMO test (KMO = .60) and Bartlett's test ($p < .001$). The analysis indicated a two-factor solution (see Table 3). Factor 1 (inhibition) included the variables inhibition speed and inhibition accuracy and factor 2 (updating) included the variables updating speed and updating accuracy. Factor scores (M = 0; SD = 1) were computed; both factors were not distributed normally (Kolmogorov-Smirnov test: $p < .05$).

Table 3. Rotated factor loadings for inhibition and updating variables

Variable	Component	
	1	2
Inhibition speed	.81	.29
Inhibition accuracy	.90	.01
Updating speed	.21	.82
Updating accuracy	.05	.86

2.3 Procedure

The data were collected in two periods: In the first period (March–April 2021), data were collected from 30 participants, and in the second period (June–September 2022) from another 33 participants. Each assessment lasted approximately 60 minutes and took place in a one-on-one setting, using an HP laptop (15-inch screen size: normed German QWERTZ keyboard). After completing the self-report questionnaire, the sentence copying task, and the EF tests, participants watched the video of the bubble gun and had as much time as needed to write the explanatory text in Microsoft Office Word, using a keyboard. This modality was chosen because adults today write texts more often with a computer keyboard than with pen and paper (Bouriga & Olive, 2021; see also Participants).

3. Results

3.1 Descriptive statistics

Participants' mean scores on the different assessments were evaluated and the variables were tested for normal distribution with the Kolmogorov-Smirnov test. Table 4 shows the descriptive statistics for all variables. All variables of the text product (except for cohesion) and the variables inhibition speed and updating speed were distributed normally. There were no missing data.

Table 4. Descriptive statistics

Variable	Measure	Mean	SD	Range
Text product				
Perceived text quality	Grade score	57.28	26.45	1.12–96.65
Information content	Number of given necessary information	8.44	2.83	1–14 (14)[a]
Cohesion	Proportion of local cohesion	0.78	0.16	0–1
Lexical diversity	Measure of textual lexical diversity	59.33	16.92	30.04–93.71
Appropriate words	Proportion of appropriate words	0.95	0.03	0.88–1
Syntactic complexity	Number of words per clause	8.67	1.19	6.18–11.73
Executive functions				
Inhibition speed	Mean reaction time	529.73	81.21	395.92–807.67
Inhibition accuracy	Number of correct responses	23.68	0.93	20–24 (24)[a]
Updating speed	Mean reaction time	597.99	118.06	358.80–845.67
Updating accuracy	Number of correct responses	13.30	2.48	3–15 (15)[a]
Shifting[i]	Mean processing time Part B–Part A	27.10	12.65	7.32–75.14

Note. [a]Theoretical maximum. [i]For further analyses, this score was inverted so that higher values corresponded to better abilities. SD = standard deviation.

3.2 Relationships between text features and perceived text quality

Bivariate Pearson or Spearman correlations were conducted to examine the relationships between the text product variables. To control for the influence of other text features, partial correlations were also performed (see Table 5).

Table 5. Correlations within the text product

	1	2	3	4	5
1. Perceived text quality					
2. Information content	.461**				
controlled	.318*				
3. Cohesion	.401**	.340**			
controlled	.299*	.342**			
4. Lexical diversity	.008	.056	-.008		
controlled	.012	.131	-.040		
5. Appropriate words	.246	.357**	.075	-.168	
controlled	.118	.374**	-.063	-.203	
6. Syntactic complexity	.010	.089	-.011	.020	-.031
controlled	-.015	.116	-.048	.000	-.069

Note. *p < .05; **p < .01.

Analyses revealed moderate positive correlations between pTQ and information content and between pTQ and cohesion, as well as between information content and cohesion and information content and appropriate words. When all other text features were included as control variables, these correlations remained significant. In the next step, a multiple linear regression analysis (using the enter method) was conducted to identify which text features predict pTQ (see Table 6).

Table 6. The influence of text features on perceived text quality

	Perceived text quality
	β
Information content	.327*
Cohesion	.282*
Lexical diversity	.010
Appropriate words	.110
Syntactic complexity	-.012

Note. *p < .05.

The model was significant, and text features accounted for 29% of the variance in pTQ. Most of the variance was explained by information content and cohesion.

3.3 The influence of executive functions on text features and perceived text quality

To determine the influence of EFs on text features and pTQ, multiple linear regression analyses (using the enter method) were conducted (see Table 7). The

regression models predicting information content and cohesion were significant. EFs explained 19.2% respectively 18.8% of the variance in information content and cohesion, although only inhibition was a significant predictor in both models. All other models did have nonsignificant predictors, although updating had a significant influence on the use of appropriate words and shifting accounted for a significant amount of variance in pTQ.

Table 7. The influence of executive functions on text features and perceived text quality

	Information content	Cohesion	Lexical diversity	Appropriate words	Syntactic complexity	Perceived text quality
	β	β	β	β	β	β
Inhibition	.406**	.347**	.092	.190	-.060	.020
Updating	.135	.057	.060	.275*	.150	.020
Shifting	.036	.192	.103	-.076	.013	.323*
R	.192**	.188**	.028	.105	.027	.110

Note. *p < .05; **p < .01.

4. Discussion

In this exploratory study we examined the complex construct of writing competence in adult, skilled writers. We assumed that the pTQ of explanatory texts would be influenced by interrelated text features especially important for this genre and that the three core EFs would influence pTQ and the relevant text features differentially in a (hitherto unexamined) way.

The results confirmed our assumptions in parts: We found that the pTQ of explanatory texts by skilled writers was influenced above all by information content and cohesion, while the use of linguistic devices did not play a decisive role. These relevant text features were both influenced by inhibition, while updating affected the use of appropriate words, and shifting had a direct impact on pTQ.

4.1 The influence of text features on perceived text quality

More specifically, 29% of the variance in pTQ could be attributed to the text features analyzed in this study, with *information content* and *cohesion* showing significant influences. This is consistent with our hypothesis and previous studies, suggesting that both information content and its organization are key factors in how raters perceive and evaluate text quality (e.g., Crossley & McNamara, 2010, 2016; Grabowski et al., 2014; MacArthur et al., 2019). Additionally, as expected, we found a significant correlation between the relevant text features. This supports the idea that a sufficient amount of information is necessary to create cohesion, yet at the

same time, the use of text structures and elements to create cohesion can facilitate the presentation of information in a text (Hennes, 2020).

However, no significant influence of *linguistic devices* on pTQ was observed in this study. Regarding *lexical diversity*, the findings align with previous studies that showed an influence on pTQ in other genres, but not in explanatory texts (Olinghouse & Wilson, 2013). In contrast, the results regarding *syntactic complexity* contradicted our hypothesis and other research (Beers & Nagy, 2009). In terms of the use of *appropriate words*, contrary to our assumptions, this study found no influence on pTQ, but a significant correlation with information content. This indicates that necessary information is only recognized as such if it is conveyed with the appropriate words. At the same time, appropriate words can only be used if the relevant information is provided. Thus, the use of appropriate words seems to be relevant for pTQ via the connection with information content.

Overall, the findings of this study support the idea that the density and connection of relevant information jointly contribute considerably to the comprehensibility of a text (Hennes, 2020). These factors are more important than using diverse and complex vocabulary or syntax, which might even hinder comprehension (McNamara et al., 2010). This aligns with the genre's purpose: to inform the reader through logically connected, complete information (Berman & Nir-Sagiv, 2007). The findings emphasize the importance of macro-level aspects of a text over micro-level aspects in explanatory texts (Donovan & Smolkin, 2006).

4.2 The influence of executive functions on text features and perceived text quality

This study shows that the three core EFs are relevant for different aspects of writing. Considering *inhibition*, we found, in line with our hypothesis, significant influences on the text features relevant for pTQ: information content and cohesion. Similar findings in previous studies with novice writers (Drijbooms et al., 2015; Zahra et al., 2023) suggested that inhibition helps suppress irrelevant information and inappropriate (lexical or syntactical) cohesive elements. It can thus facilitate the selection of relevant information, words, and phrase structures in both novice and skilled writers, contributing to the transformation of relevant ideas into a coherent, functional text (Hooper et al., 2021; Olive, 2021).

However, contrary to our assumptions and previous studies with novice writers (Altemeier et al., 2008; Rocha et al., 2022), we did not find a significant influence of inhibition on pTQ. Perhaps inhibition is not directly related to pTQ but is particularly relevant through its influence on information content and cohesion, which in turn impact pTQ. Another explanation could be ceiling effects in the measurement of inhibition accuracy. The participants consistently achieved very good results in this measure, which led to limited variability in the data. This low variance could have obscured existing direct effects of inhibition on pTQ.

Regarding *updating*, this study found no significant influences on information content, cohesion, and pTQ. This result contradicts our hypothesis and previous study results with novice writers (Bourke & Adams, 2003; Connelly et al., 2012; Drijbooms et al., 2015). It also contrasts with the assumption that updating plays an important role in revising plans and information and is particularly necessary for the writing strategy of *knowledge transforming* in skilled writers (Bereiter & Scardamalia, 1987; Ruffini et al., 2024).
We glimpsed several possible reasons for the lack of significant influences in this study.

First, the relationship between updating and writing changes with age (Ruffini et al., 2024), and updating might be less crucial for skilled writers. Second, the selected writing task might have required less updating because the video provided all relevant information and a good structure that the participants could apply without making any adaptations. Third, the participants' very good results for updating accuracy could have led to a masking of the actual influences of updating on pTQ and text features. Finally, the absence of significant relationships may also be due to the specific method used to assess updating. In this study, we employed the N-back task, which is considered to more precisely isolate updating compared to digit span tasks. The latter have been frequently used in previous research (e.g. Bourke & Adams, 2003; Connelly et al., 2012) but tend to conflate passive short-term storage with active updating operations (Owen et al., 2005; Ruffini et al., 2024). This discrepancy may suggest that short-term memory capacity – rather than updating per se – plays a more substantial role in writing performance.

However, this study did show that updating significantly influenced the use of appropriate words for skilled writers, indicating that updating plays an important role in retrieving appropriate words from long-term memory and adapting text to a reader's needs (Kellogg, 2008; Olive, 2021). Nevertheless, further studies are needed to gain a more comprehensive understanding of the role of updating in the writing process of skilled writers.

In line with our expectations and studies that identified *shifting* as a significant predictor of pTQ in novice writers (Altemeier et al., 2008; Rocha et al., 2022), we found a significant effect of shifting on pTQ in skilled writers. This supports the assumption that shifting enables skilled writers to switch between relevant information, different ideas, and activities of the writing process, managing writing processes simultaneously or adapting flexibly to writing problems or new ideas (Hooper et al., 2021; Olive, 2021). This may also explain why shifting directly impacted pTQ instead of influencing individual relevant text features, as we had assumed. However, it should be noted that commonly used tests for measuring shifting, such as the TMT, often do not assess shifting in isolation but also involve other EF components (St Clair-Thompson & Wen, 2021). This may be another reason why shifting appears relevant for text quality.

In contrast to studies with novice writers (e.g., Balioussis et al., 2012; Drijbooms et al., 2015), we found no evidence that the three core EFs influenced lexical diversity or syntactic complexity in skilled writers. This suggests that EFs may not be necessary for these features, as these features did not seem to affect pTQ in this specific writing task and therefore appeared to be less relevant or cognitively demanding. Overall, this study indicates that the core EFs play a role in the writing of skilled writers, with inhibition, updating, and shifting being associated with different aspects of (explanatory) texts. This supports the view that writing high-quality (explanatory) texts is a complex task with various (direct and indirect) pathways to pTQ not only for novice (Ruffini et al., 2024) but also for skilled writers (Olive, 2021). It thus provides further support for the assumptions of theoretical writing models being applicable to skilled writers as well (Berninger & Winn, 2006; Kim & Graham, 2022).

4.3 Limitations and future research

The limitations of this study suggest possible directions for future research in the area of skilled writing. First, our sample size was relatively small, and the recruitment method and financial compensation used may have led to biases. Second, pTQ and individual text features were measured with a single writing sample, but since the relationships between pTQ, individual text features, and EFs heavily depend on the genre and the writing task (Beers & Nagy, 2009; Kim & Graham, 2022), the results of this study only refer to explanatory texts and cannot be easily transferred to other genres or tasks. Thus, future studies might fruitfully test these relationships in skilled writers with multiple writing samples within one genre as well as in other genres or with different prompts. Additionally, the selection of the measurement methods for inhibition and updating should be reconsidered. Perhaps adapting the tasks and increasing the level of difficulty could lead to more variance in the data.

Finally, this study does not claim to be complete with regard to the selection of variables. It would be interesting to extend analyses to additional higher-level skills, such as knowledge of text structures (Hennes, 2020). Also, more complex higher-level EFs could be taken into account, as these functions could help writers compose organized, cohesive, and understandable texts (Hooper et al., 2021).

4.4 Conclusion

The findings of the current study contribute to a further understanding of the complex construct of writing competence in adult, skilled writers. The results emphasize the importance of selecting and combining relevant information when writing explanatory texts and show that the three core EFs are crucial for writing not only in novice and developing writers but also in skilled writers across the lifespan — albeit in different ways.

The findings are a first step toward further research that hopefully investigates these influences in more detail (also with longitudinal designs) to gain a better understanding of how the role of EFs for writing changes with age and expertise. Follow-up studies could subsequently open up new avenues for evidence-based diagnostics and interventions to improve the writing of adults with writing difficulties, while considering the role of EFs.

Author note

The data associated with this study is available on the Open Science Framework at https://osf.io/cxkeq/?view_only=3e0dda2ec84c495c81fea11de20e0572

References

Altemeier, L. E., Abbott, R. D., & Berninger, V. W. (2008). Executive functions for reading and writing in typical literacy development and dyslexia. *Journal of Clinical and Experimental Neuropsychology*, *30*(5), 588–606. https://doi.org/10.1080/13803390701562818

Balioussis, C., Pascual-Leone, J., & & Johnson, J. (2012). Fluency and complexity in children's writing: The role of mental attention and executive function. *Rivista Di Psicolinguistica Applicata*, *7*(3), 33–46.

Beauvais, C., Olive, T., & Passerault, J.-M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *Journal of Educational Psychology*, *103*(2), 415–428. https://doi.org/10.1037/a0022545

Becker-Mrotzek, M., Grabowski, J., Jost, J., & Linnemann, M. (2014). Adressatenorientierung und Kohärenzherstellung im Text. Zum Zusammenhang kognitiver und sprachlicher realisierter Teilkompetenzen von Schreibkompetenz. *Didaktik Deutsch: Halbjahresschrift Für Die Didaktik Der Deutschen Sprache Und Literatur*, *19*(37), 21–43. https://www.pedocs.de/volltexte/2019/17151/pdf/didaktik_deutsch_2014_37_becker

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing*, *22*(2), 185–200. https://doi.org/10.1007/s11145-007-9107-5

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Routledge. https://doi.org/10.4324/9780203812310

Berman, R., & Nir-Sagiv, B. (2007). Comparing Narrative and Expository Text Construction Across Adolescence: A Developmental Paradox. *Discourse Processes*, *43*(2), 79–120. https://doi.org/10.1080/01638530709336894

Berninger, V. W., & Winn, W. D. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). Guilford Press.

Bouriga, S., & Olive, T. (2021). Is typewriting more resources-demanding than handwriting in undergraduate students? *Reading and Writing*, *34*(9), 2227–2255. https://doi.org/10.1007/s11145-021-10137-6

Bourke, L., & Adams, A. M. (2003). The relationship between working memory and early writing assessed at the word, sentence and text level. *Educational and Child Psychology*, *20*(3), 19–36. https://doi.org/10.53841/bpsecp.2003.20.3.19

Connelly, V., Dockrell, J. E., Walter, K., & Critten, S. (2012). Predicting the Quality of Composition and Written Language Bursts From Oral Language, Spelling, and Handwriting Skills in Children With and Without Specific Language Impairment. *Written Communication*, *29*(3), 278–302. https://doi.org/10.1177/0741088312451109

Cornoldi, C., & Vecchi, T. (2003). *Visuo-spatial working memory and individual differences*. *Essays in cognitive psychology*. Psychology Press. https://doi.org/10.4324/9780203641583

Crossley, S. A., & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 984–989). https://escholarship.org/uc/item/6n5908qx

Crossley, S. A., & McNamara, D. S. (2016). Say more and be more coherent: How text elaboration and cohesion can increase writing quality. *Journal of Writing Research*, *7*(3), 351–370. https://doi.org/10.17239/jowr-2016.07.03.02

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dockrell, J. E., & Connelly, V. (2021). Capturing the challenges in assessing writing. In T. Limpo & T. Olive (Eds.), *Executive functions and writing* (pp. 103–135). Oxford University Press. https://doi.org/10.1093/oso/9780198863564.003.0005

Donovan, C. A., & Smolkin, L. B. (2006). Children's understanding of genre and writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 131–143). Guilford Press.

Drijbooms, E., Groen, M. A., & Verhoeven, L. (2015). The contribution of executive functions to narrative writing in fourth grade children. *Reading and Writing*, *28*(7), 989–1011. https://doi.org/10.1007/s11145-015-9558-z

Eckes, T. (2019). Many-facet Rasch measurement. In V. Aryadoust (Ed.), *Quantitative Data Analysis for Language Assessment Volume I: Fundamental Techniques* (pp. 153–172). Routledge. https://doi.org/10.4324/9781315187815-8

Fang, Z., Cao, P., & Murray, N. (2020). Language and meaning making: Register choices in seventh- and ninth-grade students' factual writing. *Linguistics and Education*, *56*, 100798. https://doi.org/10.1016/j.linged.2020.100798

Feenstra, H. (2021). *Assessing writing ability in primary education : On the evaluation of text quality and text complexity* [Dissertation, University of Twente]. Research Information University of Twente. https://doi.org/10.3990/1.9789036537254.

Grabowski, J., Becker-Mrotzek, M., Knopp, M., Jost, J., & Weinzierl, C. (2014). Comparing and combining different approaches to the assessment of text quality. In D. Knorr, C. Heine, & J. Engberg (Eds.), *Methods in Writing Process Research* (pp. 147–165). Lang.

Hayes, J. R. (2012). Modeling and Remodeling Writing. *Written Communication*, *29*(3), 369–388. https://doi.org/10.1177/0741088312451260

Hennes, A.-K. (2020). *Schreibprodukte bewerten: die Rolle der Expertise bei der Bewertung der Textproduktionskompetenz [Evaluating writing products: the role of expertise in assessing text production skills]* [Dissertation, Universität zu Kuln]. Kölner-Universitäts-Publikationsserver. https://kups.ub.uni-koeln.de/id/eprint/11414

Hooper, S. R., Costa, L., Fernandez, E., Barker, A., Valdes, C., Catlett, S., & & Green, M. (2021). Executive functions and writing skills in children and adolescents. In T. Limpo & T. Olive (Eds.), *Executive functions and writing* (pp. 139–159). Oxford University Press. https://doi.org/10.1093/oso/9780198863564.003.0006

Jones, I., & Karadeniz, I. (2016). An alternative approach to assessing achievement. In C. Csikos, A. Rausch, & J. Szitanyi (Eds.), *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education.*

Jost, J. (2022). Schreibaufgaben zur Indikation von Schreibkompetenz [Writing tasks to indicate writing competence]. In M. Becker-Mrotzek & J. Grabowski (Eds.), *Schreibkompetenz in der Sekundarstufe: Theorie, Diagnose und Förderung* (pp. 117–132). Waxmann.

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective journal of writing research. *Journal of Writing Research*, *1*(1), 1–26. https://doi.org/10.17239/jowr-2008.01.01.1

Kim, Y.-S. G., & Graham, S. (2022). Expanding the Direct and Indirect Effects Model of Writing (DIEW): Reading–writing relations, and dynamic relations as a function of measurement/dimensions of written composition. *Journal of Educational Psychology*, *114*(2), 215–238. https://doi.org/10.1037/edu0000564

Koizumi, R., & In'nami, Y. (2012). Effects of text length on lexical diversity measures: Using short texts with less than 200 tokens. *System*, *40*(4), 554–564. https://doi.org/10.1016/j.system.2012.10.012

Kortendiek, B., Mense, A., Beaufaяs, S., Bьnnig, J., Hendrix, U., Herrmann, J., Mauer, H., & & Niegel, J. (2022). *Gender-Report 2022: Geschlechter(un)gerechtigkeit an nordrhein-westfälischen Hochschulen.* Studien Netzwerk Frauen und Geschlechterforschung.

Lesterhuis, M., Verhavert, S., Coertjens, L., Donche, V., & Maeyer, S. de. (2016). Comparative Judgement as a Promising Alternative to Score Competences. In E. Cano & G. Ion (Eds.), *Innovative practices for higher education assessment and measurement* (pp. 119–138). IGI Global. https://doi.org/10.4018/978-1-5225-0531-0.ch007

Lezak, M. (2004). *Neuropsychological assessment* (4th ed.). Oxford University Press.

MacArthur, C. A., Jennings, A., & Philippakos, Z. A. (2019). Which linguistic features predict quality of argumentative writing for college basic writers, and how do those features change with instruction? *Reading and Writing*, *32*(6), 1553–1574. https://doi.org/10.1007/s11145-018-9853-6

Martin, J., & Rothery, J. (1980). *Writing Project Report No. I. Working Papers in Linguistics.* University of Sydney: Department of Linguistics.

Mathiebe, M. (2019). Wortschatzfдhigkeiten in der Sekundarstufe I–Plдdoyer fьr eine textorientierte Perspektive [Vocabulary skills at lower secondary level - a plea for a text-orientated perspective]. *Forschung Sprache*, *7*(3), 94–106. https://www.forschung-sprache.eu/heft-archiv/jahrgang-7-2019/heft-3/

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, *42*(2), 381–392. https://doi.org/10.3758/BRM.42.2.381

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic Features of Writing Quality. *Written Communication*, *27*(1), 57–86. https://doi.org/10.1177/0741088309351547

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

NAEP. (2011). *Writing Framework for the 2011 National Assessment of Educational Progress.* U.S. Department of Education.

Olinghouse, N. G. (2008). Student- and instruction-level predictors of narrative writing in third-grade students. *Reading and Writing*, *21*(1-2), 3–26. https://doi.org/10.1007/s11145-007-9062-1

Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing*, *26*(1), 45–65. https://doi.org/10.1007/s11145-012-9392-5

Olive, T. (2021). Executive functions in skilled writers. In T. Limpo & T. Olive (Eds.), *Executive functions and writing* (pp. 207–226). Oxford University Press. https://doi.org/10.1093/oso/9780198863564.003.0009

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 46–59. https://doi.org/10.1002/hbm.20131

Philippek, J., Kreutz, R. M., Hennes, A. K., Schmidt, B. M., & Schabmann, A. (2025). The contributions of executive functions, transcription skills and text-specific skills to text quality in narratives. Reading and Writing, 38(3), 651-670. https://doi.org/10.1007/s11145-024-10528-5

Rocha, R. S., Castro, S. L., & Limpo, T. (2022). The role of transcription and executive functions in writing: A longitudinal study in the transition from primary to intermediate Grades. *Reading and Writing*, *35*(8), 1911–1932. https://doi.org/10.1007/s11145-022-10256-8

Ruffini, C., Osmani, F., Martini, C., Giera, W.-K., & Pecini, C. (2024). The relationship between executive functions and writing in children: A systematic review. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *30*(1), 105–163. https://doi.org/10.1080/09297049.2023.2170998

St Clair-Thompson, H., & Wen, Y. (2021). Assessment of Executive Functions in Children. In T. Limpo & T. Olive (Eds.), *Executive functions and writing* (pp. 79–102). Oxford University Press. https://doi.org/10.1093/oso/9780198863564.003.0004

Struthers, L., Lapadat, J. C., & MacMillan, P. D. (2013). Assessing cohesion in children's writing: Development of a checklist. *Assessing Writing*, *18*(3), 187–201. https://doi.org/10.1016/j.asw.2013.05.001

Verhavert, S., Bouwer, R., Donche, V., & Maeyer, S. de (2019). A meta-analysis on the reliability of comparative judgement. *Assessment in Education: Principles, Policy & Practice*, *26*(5), 541–562. https://doi.org/10.1080/0969594X.2019.1602027

Vieira, A. I., Magalhães, S., & Limpo, T. (2023). Relating transcription, executive functions and text quality in Grades 2–3: A cross-lagged panel analysis. *British Journal of Educational Psychology*, *93*(2), 482–499. https://doi.org/10.1111/bjep.12570

Zahra, R. M., Sumiyadi, S., Cahyani, I., & Sastromiharjo, A. (2023). Role of Executive Functions in Improving Students' Narrative Text Writing Ability. *International Journal of Learning, Teaching and Educational Research*, *22*(6), 694–710. https://doi.org/10.26803/ijlter.22.6.35

Zimmermann, P., & Fimm, B. (2009). *Testbatterie zur Aufmerksamkeitsprüfung-Version 2.2 [Test battery for the attention test version 2.2]*. Psytest.

Appendix

Catalogue of criteria for determining the information content of texts
Function of the bubble gun
Description of the construction of the bubble gun
Soapy water is in a container
The container is screwed onto the bubble gun
A tube extends into the container
The tube leads to the opening of the gun
The pump is activated by triggering the gun
The pump is battery-operated
The soapy water is conveyed/ pumped through the tube
The ventilator is activated by triggering the gun
The ventilator is required for air supply
The soapy water is transported through/ to the blowing circle
Soap bubbles are created at the opening of the gun
The gun must be activated several times to create soap bubbles