The role of executive functions and transcription skills in writing: a cross-sectional study across 7 years of schooling

Naymé Salas[1] · **Sara Silvente[1]**

Abstract

Findings around the cognitive resources needed to compose text have helped shape current models of writing. Some of these models predict that text generation is constrained by two groups of skills: transcription (i.e., spelling and handwriting) and executive functions (EFs). While the constraining role of transcription on text generation is robust, the relationship between writing and EFs is more scarce. Some studies suggest that the impact of EFs on writing development is not only direct, but also indirect, through transcription skills. However, few studies have analyzed these effects over a sufficiently wide developmental period. In this paper, we examined transcription and EF skills in a large sample of beginner (G2), intermediate (G4), and upper-intermediate (G8) children ($N = 1337$). Each participant produced a narrative and an opinion essay. In addition, we collected measures of the low-level EFs of inhibition and updating of working memory, as well as measures of handwriting fluency and spelling accuracy. Results showed that EFs impacted text generation directly and indirectly via transcription skills, especially via handwriting. Transcription skills constrained text generation across grade levels and its weight was similar from the youngest through to the oldest age group. We conclude that EFs support low-level writing skills, as well as key processes involved in children's text composing (i.e., knowledge-telling processes). Relevant educational implications are discussed.

Keywords Writing development · Executive functions · Transcription · Text generation · Catalan

✉ Naymé Salas
nayme.salas@uab.cat

[1] Dept. de Didàctica de la Llengua i la Literatura, i de les Ciències Socials, Facultat de Ciències de l'Educació, Universitat Autònoma de Barcelona, Campus de Bellaterra, Edifici G-5, 08193 Cerdanyola del Vallès, Barcelona, Spain

Introduction

Research on writing is typically concerned with determining the cognitive under-pinnings that support text composition (Berninger et al., 1992, 1994; Graham, Berninger, Abbott, Abbott, & Whitaker, 1997; Kim, Al Otaiba, Folsom, Greulich, & Puranik, 2014). This is particularly relevant in writing development, as a solid understanding of the cognitive resources needed to write effectively should inform the teaching of writing and the identification and remediation of writing difficulties. The findings around the cognitive resources needed for writing have helped shape current models of writing (e.g., Berninger & Winn, 2006; Hayes, 2012). The not-so-simple view of writing, for example, predicts that text generation at the word, sentence, and text level is constrained by two groups of skills: transcription (i.e., spelling and handwriting) and executive functions (EFs, e.g., planning and revision, Berninger & Amtmann, 2003). When transcription skills are not automatized, children need to devote most or all of their cognitive effort to spelling and handwriting, leaving little resources available for other processes of writing, thus limiting substantially the amount of text they can generate. While the constraining role of transcription on text generation is robust (e.g., Graham et al., 1997; Juel, Griffith, & Gough, 1986; Lerkkanen, Rasku-Puttonen, Aunola, & Nurmi, 2004), few studies have analyzed a wide-enough range of developmental stages to determine whether these constraints operate mainly in the earlier grades (e.g., Juel et al., 1986; Kim, Al Otaiba, Folsom, & Greulich, 2013), or whether they extend to the entire elementary school period (e.g., Abbott, Berninger, & Fayol, 2010; Limpo & Alves, 2013a).

EFs also seem to impact on text generation (Altemeier, Jones, Abbott, & Berninger 2006; Altemeier, Abbott, & Berninger, 2008; Berninger, Abbott, Cook, & Nagy, 2017; Berninger et al., 2010; Drijbooms, Groen, & Verhoeven, 2015; Hooper, Swartz, Wakely, De Kruif, & Montgomery, 2002). There is evidence that individual variations in the level of sophistication of planning and revising strategies or in the use of a range of self-regulation strategies are positively related with text generation (Graham & Perin, 2007; Limpo & Alves, 2013b). Contrary to transcription, some EFs, such as planning or revising, will not be fully automatized (McCutchen, 2000), so their overall influence on text generation may be different in nature. Other EFs (e.g., inhibition) have also been found to contribute to levels of text generation (Drijbooms et al., 2015). Similarly to transcription, we know very little about the developmental stages at which EFs constraints operate. Moreover, few studies have looked at the interplay that may exist between transcription and EFs, and how they would impact text generation across the development of writing. The goal of this paper was, thus, to ascertain the developmental stages at which transcription and EF constraints on text generation predominate. A second aim was to understand whether the relationship between transcription, EFs, and text generation is the same throughout an ample developmental period or whether it varies across the development of writing.

Research on how transcription constraints text generation

A productive line of research has provided evidence corroborating the limiting role of transcription skills in writing (e.g., Abbott et al., 2010; Babayigit & Stainthorp, 2010; Connelly, Dockrell, Walter, & Critten, 2012; Juel et al., 1986; Kim et al., 2011; Puranik, Lombardino, & Altmann, 2008; Puranik & Al Otaiba, 2012). Spelling and handwriting have been argued to constrain text generation particularly in the earlier stages of learning to write, when these skills have not been automatized (e.g., Berninger et al., 1992; Juel et al., 1986; Kent, Wanzek, Petscher, Al Otaiba, & Kim, 2014; Kim et al., 2011). However, some studies have found spelling and handwriting to exert a protracted influence in writing (e.g., Alves & Limpo, 2015; Limpo & Alves, 2013a). Wagner et al. (2011), for example, found that handwriting had a stronger correlation with macro-structural quality in the texts produced by a group of 4th graders, than in a group of 1st graders. Finally, transcription skills, particularly spelling, vary in their complexity depending on the orthography involved. A longitudinal study that compared the role of spelling skills in early (1st and 2nd grade) text generation in English—which has a very inconsistent and complex orthography—with Spanish—which has a very consistent and simpler orthography—found that spelling constrained text generation in English for a longer period than in Spanish, where it explained little variance after the end of 1st grade (Salas & Caravolas, submitted). To sum up, while it seems that transcription skills pose limits to text generation in early writing, it is not clear until when these limitations operate. Furthermore, languages other than English, whose transcription skills may be mastered earlier, are rarely examined (e.g., Caravolas & Bruck, 1993). In this paper we analyzed texts written in Catalan, a Romance language with a semi-consistent orthography, in three cohorts of children attending 2nd and 4th grade of Primary school, and 2nd year of Secondary education (i.e., 8th grade).

Research on how EFs constrain text generation

A distinction has been made in the literature between low-level and high-level EFs (Altemeier et al. 2008; Drijbooms et al., 2015). High-level EFs are those often described in most writing models, such as planning, revising, or self-regulation skills (Berninger & Winn, 2006; Hayes, 2012; Hayes & Flower 1980). Writers who apply efficient planning and revising strategies often generate more text and their written expression is of a higher quality (Berninger, Whitaker, Feng, Swanson, & Abbott, 1996; Limpo & Alves, 2013a). Expert writers usually report using sophisticated planning techniques and revising several versions of their texts (Bryson, Bereiter, Scardamalia, & Joram, 1991; Hayes & Flower 1980). Another high-level EF is self-regulation, which is a kind of goal-oriented behaviour involving the endorsement and monitoring of some standards of thought (Hofmann, Schmeichel, & Baddeley, 2012). Writers with more developed self-regulation strategies often compose more and better-quality texts (Graham & Perin, 2007; Graham, McKeown, Kiuhara, & Harris, 2012). These high-level EFs are of a complex nature (Hayes & Flower,

1980), whose internal architecture has been argued to be composed of low-level EFs (e.g., Drijbooms et al., 2015). Typically, three low-level EFs are considered: inhibition, working memory (WM) and cognitive flexibility (Diamond, 2013; Miyake et al., 2000). Inhibition refers to the ability to either suppress attention to non-relevant stimuli or prepotent, automatic responses. WM involves actively holding verbal or visuospatial information in mind while performing mental operations (e.g., updating of information) upon them. Finally, shifting or cognitive flexibility refers to the abilities consisting of (1) fluently generating and using ideas that, in turn, lead to fluent responses, and (2) shifting between mental mindsets (Diamond, 2013; Hofmann et al., 2012; Miyake et al., 2000; Vandenbroucke, Verschueren, & Baeyens, 2017). These core EFs would cooperatively underlie more complex or high-level EFs such as planning or revising (Vanderberg & Swanson, 2007). Arguably, these low-level skills may provide a detailed insight into the cognitive structures behind written composition at different stages in development. Moreover, research with young writers has found that they rarely plan or revise texts and that their planning and revising skills do not translate into longer or better texts (Berninger et al., 2010; Limpo & Alves, 2013a); therefore, these domain-specific high-level processes of writing may not be useful for comparisons over a wide developmental span. In contrast, domain-general, low-level EFs like inhibition and updating of WM have been found to be operative at these ages and before, while they continue to develop in the later school years (Best, Miller, & Jones, 2009; Diamond, 2013). Moreover, they have been reported to contribute to explaining variance in text generation in grade 1 children (e.g., Berninger et al., 2010), as well as in older adults (Hoskyn & Swanson, 2003).

The importance of low-level EFs to studies on writing development stems from our current understanding that writing occurs in a working memory environment (Berninger & Winn, 2006). This means that the series of knowledge sources, processes, and skills involved in producing text must compete for a limited pool of cognitive resources (Kellogg, 2001; Olive & Kellogg, 2002). When a particular skill (e.g., spelling) is immature, it limits the amount of items that a writer may hold in her center of attention at any given point (e.g., the letters or words that the writer wants to represent). Conversely, if processing is fluent or automatic, it releases resources, so that more items can be consciously hold and manipulated (e.g., more words, letters, or ideas at any one time). This tradeoff between efficient or fluent processing and storage capacity accommodates the robust evidence of the increase in average burst length as a function of age (Ailhaud, Chenu, & Jisa, 2016; Alves & Limpo, 2015). It also points to the WM paradox (McCutchen, 2000): on the one hand, novice writers tax their available cognitive resources due to disfluent (immature) transcription, which, in turn, compromises text generation; but more experienced writers, on the other hand, who have automatized transcription, also tax their available cognitive resources with other concerns, such as goal setting, interlocutor awareness, genre well-formedness, and a myriad of other rhetorical problems (Bryson et al., 1991). This happens because the representation of the writing task differs between novice and more experienced writers: While the former adopt a knowledge-telling strategy that essentially involves generating content that gets transcribed, the latter adopt a knowledge-transforming strategy that involves complex content and

rhetorical problems (Bereiter & Scardamalia, 1987). However, while less experienced writers would be overloading WM capacity with transcription-related tasks, more experienced writers progressively incorporate quickly-to-access solutions to resolve frequent rhetorical issues. In this way, the experienced writer may not only make optimal use of her memory capacity, but may even surpass WM limitations (McCutchen, 2000). Put differently, while novice writers are chiefly occupied with letter retrieval and formation (i.e., handwriting), or sounding out parts of words to match the content as it is being generated (i.e., spelling), expert writers are chiefly occupied with finding solutions to various rhetorical issues. These differences in the writing process between novice and more experienced writers should lead to a constant, but essentially different, use of EFs to support text generation. Specifically, we would expect that EFs in novice writers support mainly transcription processes which, in turn, constrain text generation processes (i.e., an indirect effect); for more experienced writers, for whom transcription would be fairly automatized, EFs would have a direct impact on text generation, as they would be supporting more complex rhetorical problem-solving processes. Consequently, in this paper we examined direct and indirect ways in which EFs support text generation at the word and sentence levels.

There is conflicting evidence about the developmental stage(s) at which low-level EF constraints predominate. In a seminal study of the role of memory resources in writing, Swanson and Berninger (1996) found that it was STM that was primarily related to transcription skills, while WM, which involves short-term storage, like STM, but also processing (i.e., conscious manipulation), was mostly involved in higher-level writing domains, like text generation (see also Puranik, 2006, for almost identical results). More recently, however, Berninger et al. (2010) found that WM was, actually, involved in writing before grade 4 and amended their initial claim (p. 188). Other studies have examined the nature of the impact of low-level EFs on writing. For example, Drijbooms et al. (2015) conducted a study with grade 4 Dutch children and tested the direct and indirect influence, through transcription skills, of a host of EF measures on writing at the word, sentence, and text levels in narrative writing. They found that a factor of inhibition and another of updating of WM influenced the number of words produced by the children (word level); the mean number of words per T-Unit, a measure of syntactic complexity (sentence level); and a measure of story content quality (text level). The impact of EFs on the word level was both direct and indirect, while influence on the sentence and text level was indirect. All indirect effects were mediated by handwriting, while spelling did not exert much influence on any level. Their findings highlight the vital role for mid-primary school children of EFs in text productivity, which is a fundamental trait of writing development and proxies several other text quality indicators (Berman & Nir-Sagiv, 2009; Berman & Verhoeven, 2002). It also means that, at least until grade 4, transcription skills still mediate and constrain text generation, but that they do so differently as a function of the level of language under investigation. Finally, it points to inhibition and updating of WM as key low-level EFs involved in written composition processes.

In conclusion, both low- and high-level EFs are involved in writing across development. Examining low-level EFs, particularly inhibition and updating, might

provide a detailed account of the recruitment of cognitive skills that is necessary for generating text and may be particularly relevant in studies including a wide developmental span.[1] For these reasons, we included measures of inhibition and updating of WM and examined their relationship with text generation in a large sample of children at different stages of writing development.

The present study

This study assessed the impact of transcription skills and of specific core, low-level executive functions on text generation skills: inhibition and updating of WM. Because previous research has suggested that the impact of EFs on text generation might depend on the age or level of expertise of the writers (e.g., Altemeier et al. 2008), we analyzed three distinct stages of writing development: beginner (2nd graders, G2), low-intermediate (4th graders, G4), and upper-intermediate (8th graders, G8). Other studies have indicated that the precise manner in which EFs affect text generation—i.e., directly or indirectly, via transcription (Drijbooms et al., 2015), may depend on the level of language under study; we thus analyzed text generation at the word and sentence (i.e., clause) levels. Finally, previous research has found that the discourse genre to be composed could affect whether and how EFs influence text generation (Altemeier et al. 2006). For this reason, we analyzed samples of narrative and opinion-essay writing from all participating children.

Our first research question addresses the role of transcription in text generation. Our working hypothesis was that text generation by children in the youngest group (G2) would be highly constrained by transcription skills, as they will not have been fully automatized and, thus, they should deplete WM resources. In contrast, we expect that, for older children, the effect of transcription would decrease progressively. Our second research question addressed the role of EF in text generation. In line with previous studies (e.g., Drijbooms et al., 2015) we expected updating and inhibition to account for variance in text generation at the word and sentence levels. We expected an indirect effect, especially in the beginner writers and also in the intermediate writers, though to a lesser extent, on both levels of text generation, because EFs might support transcription skills which, in turn, would support text generation. A direct effect was expected especially for older participants (i.e., G8), as EFs should be supporting writing processes other than transcription. Moreover, we hypothesized that composing narrative texts would be, in general, less cognitively demanding than composing opinion essays, given children's higher familiarity with narrative discourse (e.g., Kellogg, 1994, 2001; Siegler, 1996). In line with previous studies suggesting that the cognitive effort for composing narratives plateaus around G5, while the cognitive effort for composing argumentative texts decreases

---

[1] There is an ongoing debate over whether the three core low-level EFs, inhibition, updating of WM, and cognitive flexibility are completely discernable throughout development. Compelling evidence has been found for cognitive flexibility to greatly overlap with other low-level EFs, particularly inhibition, in children (e.g., Drijbooms et al., 2015; van der Ven, Kroesbergen, Boom, & Leseman, 2013). Hence, we opted for targeting only inhibition and updating of WM.

Table 1 Demographic information by grade level

	Grade 2	Grade 4	Grade 8
N	486	485	366
M Age in years	7;5	9;4	13;6
(SD)	(0.52)	(0.50)	(0.53)
Sex (boys)	229	252	195
M ISEI[a] average	46.08	46.85	39.53
(SD)	(17.10)	(16.82)	(11.61)
Familiarity with Catalan[b]	48.8%	48.8%	76.2%

[a]ISEI: International Socio-economic Index (Ganzeboom et al., 1992)

[b]Familiarity with Catalan is shown here as the percentage of children in each grade level who speak Catalan at home, with either or both parents

as a function of schooling experience (Olive, Favart, Beauvais, & Beauvais 2009), we expected that the direct effect of EFs on narrative writing would increase from G2 to G4, and remain stable afterwards. In contrast, we expected our EF measures to have a larger, direct effect on the production of opinion essays in G2 and G4, than in G8.

This study thus improves on previous ones by presenting data from different grade levels, covering a wide range of writing development throughout the elementary school years. In contrast with previous studies, careful attention was paid to controlling for the socioeconomic status of the participants, as well as for sex-related influences, which are known to affect writing development (Cordeiro, Castro, & Limpo, 2018; Hajovsky et al. 2018; Hillocks, 2006). Finally, we explored writing in Catalan, a Romance language with a semi-transparent orthography, morphosyntactically similar to Spanish, which has been rarely explored, thus contributing to the analysis of writing development in languages other than English.

Method

Participants

Participants were 1337 children (676 boys) attending G2, G4, and G8. They were recruited from 13 public Primary schools (the 2nd and 4th graders) and 5 high-schools (the 8th graders) in the province of Barcelona (Spain). Detailed demographic information as a function of grade is reported in Table 1. The project had obtained approval by the Ethics Committee of the University. Parents were given informative letters and consent forms. In the case of the 8th graders, self-consent was also required. Information about participants' socioeconomic status (SES) was obtained through a questionnaire completed at school. The questionnaire inquired about participants' language use and about their parents' jobs and education level. Questions about the languages spoken at home intended to estimate their overall familiarity with and exposure to Catalan, which is the language of instruction at school. It

should be noted that Barcelona belongs to a region in Spain, Catalonia, where virtually everyone is, at least, bilingual (Spanish–Catalan). Familiarity with Catalan was estimated based on children's reported language use with both parents. For each parent with whom they spoke Catalan, we afforded them 1 point; if they spoke Catalan as well as another language (e.g., Spanish), we afforded them 0,50 points. Finally, if Catalan was one of three possible languages they used with either parent, we afforded them 0,33 points. The resulting variable thus oscillated between 0 (no use of Catalan with either parent) and 2 (Catalan is the main language used with both parents[2]), and was used as a covariate in subsequent analyses. Information about parents' jobs was used to estimate SES, using Ganzeboom, De Graaf, and Treiman (1992) International Socio-Economic Index (ISEI). This instrument was chosen because it was suitable to accommodate children whose families are from very diverse backgrounds and countries of origin, so it was important to ensure international validity. Furthermore, because it consists of a continuous scale (in this sample, the minimum score was 17, for a unskilled farmer, and 88 for a series of highly qualified jobs, such as dentist or surgeon), it allowed including SES as a covariate in the models.

Tasks and measures

Inhibition

The opposite-worlds task, a stroop-like task adapted to children (Strauss, Sherman, & Spreen, 2006), measures complex response inhibition, since it requires to "hold a rule in mind, respond according to this rule while inhibiting the prepotent response associated to the perceptual stimulus" (Montgomery & Koeltzow, 2010). Participants were shown a card with a circuit of 40 squares containing either the number one (1) or the number two (2), pseudorandomly distributed. In the "same-world" trial children were asked to follow the circuit naming the number they saw (e.g., "2, 2, 1, 2, 1…") as quickly as possible. Two "opposite-world" trials ensued, in which children were told to say the opposite to what they saw; that is, say 1 for number 2, and 2 for number 1. In all trials, the administrator registered the number of errors produced and the overall time to complete the trial. Trials with more than 15% errors were excluded. The score was determined by calculating the mean time (in seconds) of the two opposite-world trials. Reliability was the correlation between the two opposite-world trials, $r = .823$.

WM updating

The *Digits* subtest of the Spanish adaptation of the WISC-IV test battery (Wechsler, 2005) task measures phonological working-memory abilities in terms of storage (forward condition) and executive capacities (backwards condition). In the forward condition, children were asked to repeat sequences of numbers in the same order as they

---

[2] Monoparental families were adjusted accordingly (i.e., multiplied by 2).

had been orally presented by the administrator. The task is composed of 8 blocks, each containing two different sequences of the same length. Blocks start two digits and increase by one digit until reaching 9-digit sequences. The test was discontinued after the child failed both sequences in a block. The backwards condition followed the same procedure, except that participants were asked to repeat the sequence of numbers in reverse order. One point was afforded for each correct trial and the score of the whole task was determined by calculating the sum score of the forward and backwards conditions. The manual reports a split-half reliability of .74, but since the task was administered in Catalan and not Spanish (the language of standardization), we estimated our own reliability estimate. Chronbach's alpha was .74.

Handwriting fluency

Handwriting fluency was assessed with the alphabet writing task (Berninger et al., 1992), which requires to write down the alphabet in order and in lowercase letters. The score obtained is the number of letters that children produced in the first 15 s. The validity of this task highly depends on the alphabet being well-known so it can be retrieved from memory and automatically noted down. However, we became aware that not all of our participants had been taught the alphabet at school. For this reason, we implemented two scoring criteria, a strict one, where errors of order were penalized, and a lenient one, where they weren't. Correlation between the two scoring procedures was very high, $r = .996$. Therefore, we used the lenient scoring, as it allowed us to use the entire sample. A well-trained research assistant scored a random selection of 30% of the sample. The ICC between the two scorers was .984.

Spelling

Children were asked to write 34 words to dictation, which were presented first in isolation (e.g., *vent,* 'wind'), then, in a carrier sentence, to clarify their meaning and context of use (e.g., *Si fa vent, podrem anar a navegar.* 'If there's wind, we can go sailing"), then repeated one last time in isolation. Children were encouraged to write as they could and the task was never discontinued. Each correctly spelled word was afforded 1 point and the final score was the sum of all correct answers. Internal consistency was assessed with Chronbach's alpha, $\alpha = .919$. Because there are no standardized instruments to assess spelling in Catalan, we examined the inter-rater reliability of this bespoke test. An experienced master's student, alien to the goals of this study, scored all writing samples and the first author scored a random 25%. ICC between the two raters was .989.

Writing samples

Children produced an opinion essay and a narrative text. There were two similar possible prompts for each genre, which were pseudo-randomly assigned to each class, to ensure roughly a 50–50 distribution of the prompts. Participants were given 10 min to plan their texts and 20 min to write them. All writing samples were transcribed in a text editor and divided into clauses (following Berman & Slobin, 1994)

Table 2 Descriptive statistics and results of one-way ANOVAs of the grade level effect

Measure	2nd grade M (SD)	4th grade M (SD)	8th grade M (SD)	F
Transcription				
Handwriting	4.70 (2.21)	7.69 (3.25)	13.89 (6.24)	590.32*
Spelling (range 0–34)	7.65 (5.95)	13.10 (6.09)	19.88 (5.85)	400.38*
Executive functions				
Inhibition (opp. worlds)	53.67 (12.77)	40.95 (8.39)	31.34 (6.43)	561.48*
Updating (digits, range 0–16)	10.68 (2.32)	12.21 (2.57)	13.28 (2.42)	123.97*
Text generation				
No. words—narrative	55.03 (34.64)	103.51 (54.56)	157.75 (80.91)	301.02*
Words/clause—narrative	10.42 (6.97)	19.55 (10.54)	28.75 (16.19)	246.65*
No. words—opinion	21.49 (16.09)	44.18 (25.44)	89.20 (48.29)	463.75*
Words/clause—opinion	5.47 (3.55)	10.07 (5.40)	16.59 (8.94)	322.45*

*$p < .001$

and later transformed into CLAN files using the TEXTIN command. A semi-random sample of 100 texts was transcribed by an experienced research assistant and the second author and inter-rater reliability estimates were obtained for the total number of words, ICC = .996; and for the total number of clauses, ICC = .958. The rest of the written samples were transcribed by the same two people independently, and doubts about criteria were discussed weekly. Using the FREQ command we obtained automatic counts of words and clauses. Text generation at the word level was estimated by the total number of words in each genre; text generation at the sentence level was the average number of words per clause in each genre.

Procedure

EFs tasks and handwriting were administered individually in a single session by well-trained research assistants in a quiet room at the children's school. There were two orders of tasks administration, which were counterbalanced. The writing samples and the spelling task were administered in one-week apart, whole-class sessions at children's regular classrooms, with their teacher and at least one research assistant, who conducted the sessions, present.

Results

Descriptive statistics are shown in Table 2, together with the results of a series of one-way ANOVAs, to test the effect of grade level on all variables. Skewness and kurtosis values indicated that measures were generally normally distributed (skewness values < 3; kurtosis values < 10, Kline, 2011). The series of ANOVA tests revealed that EF, transcription, and text generation measures improved significantly with age and schooling. Post-hoc tests (Bonferroni) revealed that all comparisons

(i.e., 2nd vs. 4th grade; 4th vs. 8th grade; and 2nd vs. 8th grade) were significant. This means that children's EF abilities continued developing throughout grade levels, that handwriting skills gained fluency, and that children increased the average length of their texts in words and in the average number of words per clause, while acquiring higher levels of spelling accuracy.

Correlations between measures

Tables 3 and 4 show the bivariate correlations between all variables at G2 and G4, and at G8, respectively. Inhibition and updating of WM were associated with each other at all grade levels, with correlations in the small to moderate range. These EF skills were, in general, significantly correlated with transcription skills and with text generation at the word level, but not at the sentence level. Handwriting and spelling were related to each other at all grade levels, with the size of correlations increasing from G2 through to G8. Text generation at the word level in narrative and in opinion-essay writing were significantly correlated with each other at all grade levels, with rs ranging from .291 to .508, thus increasing with schooling experience. Text generation at the sentence level in narrative texts did not significantly correlate with the same measure in opinion essays at any grade level, but both occasionally correlated with text generation at the word level. Finally, text generation was significantly correlated with transcription at the word level, though not at the sentence level.

Path analyses and comparisons between grade levels

In order to assess the role of EFs and transcription in text generation across schooling, we ran multigroup path analyses in MPlus v. 8.0 (Muthén & Muthén, 1998). Separate models were specified for each type of dependent variable; that is, word generation at the word and sentence level. In each model, we kept text generation in narratives and opinion essays as two different dependent variables, given that children produced significantly more words and more words per clause in the narratives than in the opinion-essays (M number of words for narratives = 100.27, SD 69.42; for opinion-essays = 47.94, SD 40.60, $t(1118) = 33.50$, $p < .001$. M number of words per clause in narratives = 5.46, SD 1.36, $d = 0.92$ a large effect size; in opinion-essays = 4.50, SD 1.57, $t(1118) = 15.94$, $p < .001$, $d = 0.65$, a medium-large effect size (Sawilowsky, 2009). In order to better assess the role of the different transcription skills, and because they were correlated at all grades, separate models were conducted for spelling and handwriting as the mediator. Predictor variables included inhibition and updating. SES, sex, familiarity/exposure to Catalan, and prompt type were included as covariates for all variables in the model.[3] To obtain highly reliable estimates, a bootstrap procedure of 1000 iterations was applied.

---

[3] Prompt-type was excluded as a covariate for the words-per-clause measure in the narrative texts, when spelling was the mediator,, as it did not explain any additional variance and it worsened model fit considerably.

Table 3 Bivariate correlations at grades 2 and 4

	1	2	3	4	5	6	7	8	9	10	11
1. Handwriting	–	.16*	-.27*	.26*	.21*	.11*	.09	.01	.18*	.12*	.12*
2. Spelling	.25*	–	-.07	.32*	.13*	.10	.06	-.07	.15*	.05	.19*
3. Inhibition[a]	-.34*	-.24*	–	-.20*	-.15*	-.11*	-.08	-.05	-.06	.08	-.05
4. Updating[b]	.20*	.29*	-.28*	–	.27*	.08	.07*	-.10*	.41*	.00	.30*
5. Words-narr	.32*	.10	-.27*	.21*	–	.07	.29*	.01	.23*	.10*	.13*
6. wcl-narr	.01	.21*	-.11	.09	.14	–	-.02	.03	-.09	.02	-.06
7. Words-opinion	.23*	.10*	-.17*	.14*	.46*	<.01	–	.31*	-.01	.04	.02
8. wcl-opinion	.03	-.03	.03	-.04	.09	-.01	.31*	–	-.22*	.06	-.22*
9. SES	.15*	.17*	-.06	.33*	.19*	-.08	.16*	-.09	–	-.04	.50*
10. Sex	.07	-.01	.04	.00	.24*	-.02	.21*	.11*	.02	–	-.07
11. Catalan	.07	-.01	.04	.00	.24*	-.02	.21*	.11*	.56*	.05	–

Grade 2 coefficients are shown above the diagonal; grade 4 coefficients are shown below the diagonal. words-narr = average number of words in narrative text; wcl-na = average number of words per clause in narrative text; words-opinion = average number of words in opinion essay; wcl-op = average number of words per clause in opinion essay; Catalan = degree of exposure and familiarity with Catalan

*p < .05

[a]Opposite-worlds task

[b]Digit-span task

Table 4 Bivariate correlations at grade 8

	1	2	3	4	5	6	7	8	9	10	11
1. Handwriting	–	.43*	−.36*	.28*	.36*	.21*	.27*	.12	.00	.24*	.03
2. Spelling		–	−.34*	.48*	.29*	.16	.35*	.16*	.05	.05	.16*
3. Inhibition[a]			–	−.35*	−.27*	−.02	−.25*	−.08	−.11	.05	.04
4. Updating[b]				–	.19*	.06	.26*	.09	.10	−.09	.14*
5. Words-na					–	−.04	.51*	.14*	.14*	.35*	.19*
6. wcl-na						–	.07	.16	−.13*	−.08	.04
7. Words-op							–	.27*	.04	.29*	.13*
8. wcl-op								–	.04	.05	.03
9. SES									–	−.14*	.25*
10. Sex										–	−.06
11. Catalan											–

Grade 2 coefficients are shown above the diagonal; grade 4 coefficients are shown below the diagonal. words-narr = average number of words in narrative text; wcl-na = average number of words per clause in narrative text; words-opinion = average number of words in opinion essay; wcl-op = average number of words per clause in opinion essay; Catalan = degree of exposure and familiarity with Catalan

*$p < .05$

[a]Opposite-worlds task

[b]Digit-span task

Models in which all parameters were fixed to be equal across groups were a poor fit to the data (Chi$^2 > 109.24$, $p < .001$, RMSEA $> .059$, CFI $< .814$, TLI $< .740$, SRMR $> .122$). Therefore, we ran a series of models in which parameters for each grade level were freely estimated and, in order to ascertain developmental changes, we included planned comparisons of all direct and indirect effects of EFs on the text generation variables, and on the effect of transcription on text generation.

Impact of EFs on text generation at the word and sentence level

Figure 1 displays a model where text generation at the word level (no. of words) in narrative and opinion-essay writing were the criterion variables, where updating and inhibition were the exogenous variables, and handwriting was the mediator variable. This model, which was freely estimated for each grade level, fitted the data very well, Chi^2 (44) $= 64.08$, $p = .026$, RMSEA $= .031$ (CIs 0.011–0.047), CFI $= .980$, TLI $= .952$, SRMR $= .031$. An identical model where spelling was the mediator variable was also an excellent fit to the data, Chi^2(44) $= 77.72$, $p = .001$, RMSEA $= .040$ (CIs 0.025–0.055), CFI $= .964$, TLI $= .915$, SRMR $= .034$ (Fig. 2). In these models, the direct effect of inhibition on narrative writing was statistically significant for all grade levels and for G4 and G8 in opinion-essay writing, while it increased significantly over time in both text types (Table 5). The indirect effect of inhibition on the number of words in narrative writing via handwriting was significant for G4 only, in both text types. Accordingly, the contrasts between G2 and
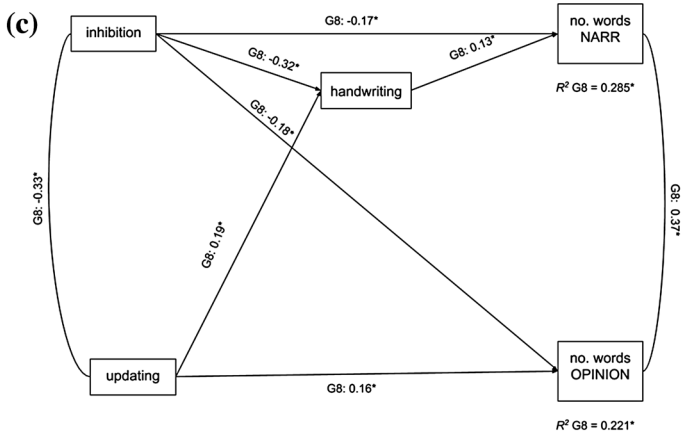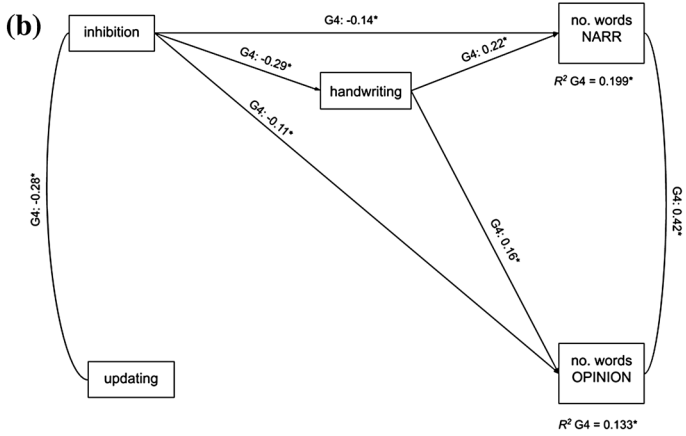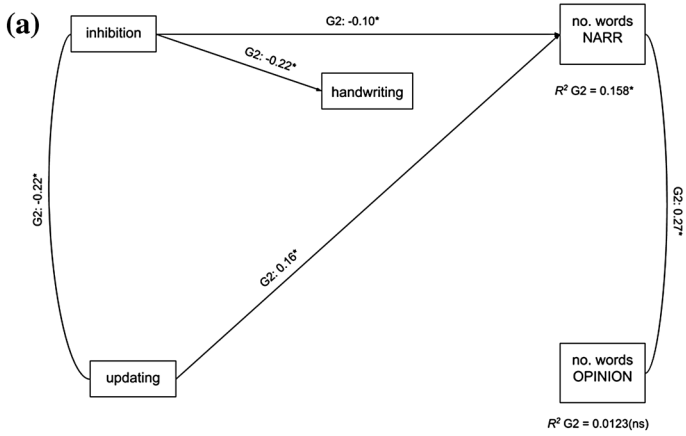
Fig. 1 Standardized coefficients of the path analysis of the influence of EF measures on text generation ▶ at the word level at **a** Grade 2, **b** Grade 4, and **c** Grade 8, with handwriting as the mediator, and controlling for the effect of SES and sex (not shown, flor clarity). *NARR* narrative text, *OPINION* opinion-essay. *$p < .05$

G4 were significant in both genres. However, we expected that the contrast between G4 and G8 would also be significant, but it was not, nor was the contrast between G2 and G8. To resolve this apparently contradictory result, we tested a model in which the indirect effect of inhibition via handwriting was constrained to be equal at all grades. This latter model fitted the data significantly more poorly than the unconstrained model, $\Delta \text{Chi}^2(6) = 32.27$, $p < .001$; $\Delta \text{CFI} = 0.028$.[4] Nonetheless, in a partial-moderation model, where only the path from handwriting to number words was constrained to be equal across age groups in both genres (see section below on transcription effects), the indirect effect of inhibition via handwriting was significant for all age groups, while it increased significantly as a function of age (Table 6). The indirect effect of inhibition via spelling was nonsignificant, except for G8 in both genres. Accordingly, the contrasts between grades showed that it increased significantly with age (Table 5).

In the model where handwriting was the mediator, updating of WM had a direct effect on the number of words in the narratives at G2 but, surprisingly, there were no significant differences between grades. For this reason, we ran an alternative model where the direct effect of updating was fixed to be equal across grades in this genre. The resulting model was a similar fit to the data, $\Delta \text{Chi}^2(2) = 0.07$, $p > .005$; $\Delta \text{CFI} = 0.001$, and it indicated that the direct effect of updating on the number of words in the narratives was significant (unstandardized estimate $= 2.27$, SE 0.60, $p < .05$). In the opinion essays, the direct effect of updating was significant at G8 only, and the contrasts between grades showed that it increased significantly from G2 and G4 to G8. In the model where spelling was the mediator, a similar situation emerged: the direct effect of updating of WM was significant at G2 in the narratives but no contrasts between grades were significant. In the case of opinion essays, the direct effect of updating was never significant but the contrast between G2 and G4 indicated a significant increase with age. To resolve these apparently conflicting results, we ran an alternative model where spelling was the mediator and the direct effect of updating on the number of words in both genres was constrained to be equal across grades. This model was a similar fit to the data as the unconstrained model, $\Delta \text{Chi}^2(4) = 2.65$, $p > .005$; $\Delta \text{CFI} = 0.002$, and it indicated that the direct effect of updating on the number of words in the narratives was significant (unstandardized estimate $= 2.27$, SE 0.60, $p < .05$), but not in the essays (unstandardized estimate $= 0.50$, SE 0.28, $p > .05$). Updating of working memory did not have an indirect effect via handwriting at any grade nor genre, but it did support the number of words indirectly via spelling at G8 in both genres. In sum, inhibition had a direct

---

[4] A ΔCFI of less that .01 between a more constrained and a less- or unconstrained model indicates that the models fit the data similarly well and is preferred over the ΔChi^2 test, particularly in models applied to large samples (Cheung & Rensvold, 2002; Meade, Johnson, & Braddy, 2008; Wang & Wang, 2012).

(a)

inhibition → no. words NARR: G2: -0.10*
inhibition → handwriting: G2: -0.22*
inhibition ↔ updating: G2: -0.22*
updating → no. words NARR: G2: 0.16*
R^2 G2 = 0.158*
no. words NARR ↔ no. words OPINION: G2: 0.27*
R^2 G2 = 0.0123(ns)

(b)

inhibition → no. words NARR: G4: -0.14*
inhibition → handwriting: G4: -0.29*
inhibition → no. words OPINION: G4: -0.11*
handwriting → no. words NARR: G4: 0.22*
handwriting → no. words OPINION: G4: 0.16*
inhibition ↔ updating: G4: -0.28*
R^2 G4 = 0.199*
no. words NARR ↔ no. words OPINION: G4: 0.42*
R^2 G4 = 0.133*

(c)

inhibition → no. words NARR: G8: -0.17*
inhibition → handwriting: G8: -0.32*
inhibition → no. words OPINION: G8: -0.18*
handwriting → no. words NARR: G8: 0.13*
updating → handwriting: G8: 0.19*
inhibition ↔ updating: G8: -0.33*
updating → no. words OPINION: G8: 0.16*
R^2 G8 = 0.285*
no. words NARR ↔ no. words OPINION: G8: 0.37*
R^2 G8 = 0.221*

Springer

and an indirect influence on the number of words across grades and in both genres, while its importance increased with grade. Updating of WM, on the other hand, had a stable direct effect on the narratives a much more limited effect on the opinion essays. Handwriting did not mediate its effects at any grade or genre, while spelling did mediate its effect at G8.

A model where narrative and opinion-essay text generation at the sentence level (words per clause) were the two criterion variables, where handwriting was the mediator variable, and updating and inhibition were the exogenous variables fitted the data well, Chi^2 (45) = 66.03, $p = .022$, RMSEA = .032 (CIs 0.012–0.047), CFI = .962, TLI = .912, SRMR = .031 (Fig. 3). Contrary to text generation at the word level, no direct or indirect effects of inhibition or updating were found at any grade level in either narrative or opinion-essay writing (Table 7). Another model, identical to the last one but where spelling was the mediator, was also a good fit to the data, Chi^2 (36) = 54.59, $p = .024$, RMSEA = .033 (CIs 0.012–0.050), CFI = .963, TLI = .908, SRMR = .031 (Fig. 4). Here, too, no direct or indirect effects of inhibition or updating were attested on either dependent variable or grade. All in all, these data suggest that EFs did not exert much influence on text generation at the sentence level, regardless of the discourse genre being produced, of the writer's experience, or of the mediator skill.

Transcription constraints in text generation

The above-mentioned models allow an examination of the role of transcription across the three developmental stages. At the word level (Table 5), the amount of text that children produced in their narratives was constrained by their handwriting skills at G4 and G8 in the narratives, and at G4 in the opinion essays. However, given that there were no significant differences in the comparisons between grades, we ran an alternative, partial-mediation model that was a similar fit to the data, $\Delta Chi2(4) = 3.64$, $p > .05$, $\Delta CFI = 0.001$, where handwriting was fixed to be equal across grades and genres. This model indicated that handwriting had a significant effect in the number of words in narrative (unstandardized estimate = 2.76, SE 0.59, $p < .05$) and in opinion essays (unstandardized estimate = 0.93, SE 0.29, $p < .05$). Contrary to handwriting, spelling did not constrain the number of words generated in the earlier grades, but it did so at G8 in both genres.

At the sentence level, neither handwriting nor spelling supported the average number of words per clause in the narrative texts or in the opinion essays. To sum up, the influence of transcription on text generation occurred predominantly at the word level. In addition, it was handwriting, rather than spelling, the transcription skill that impacted the most on the amount of text that children generated.
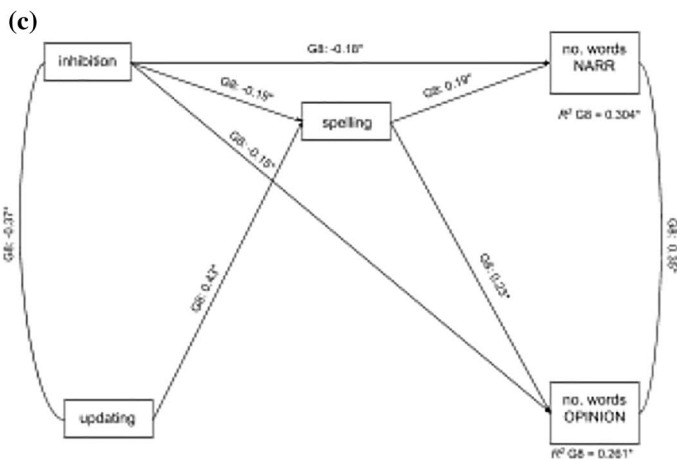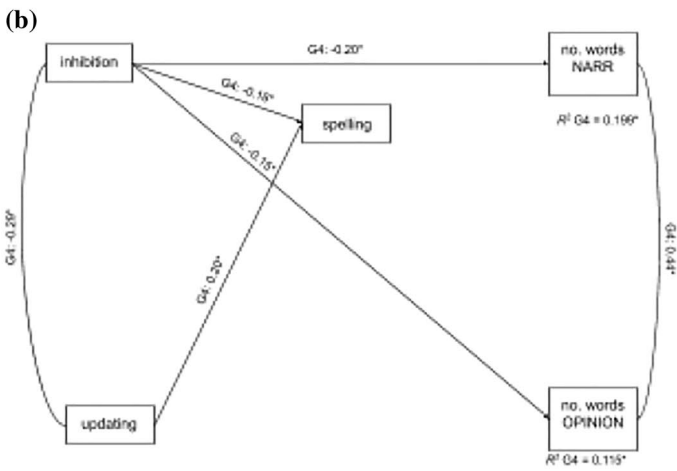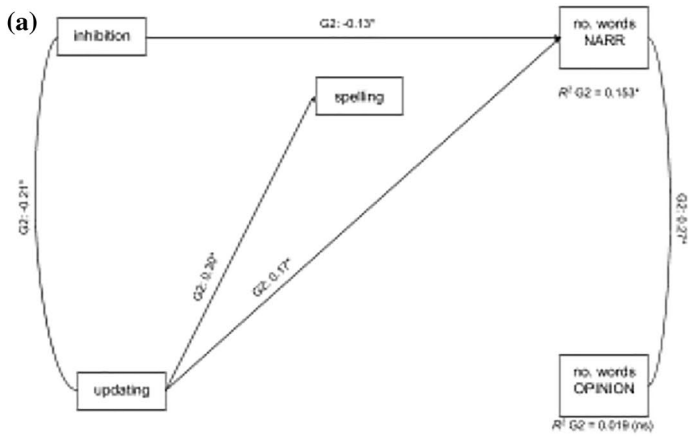
(a)

inhibition
G2: -0.13*
no. words NARR
R^2 G2 = 0.153*
spelling
G2: -0.21*
G2: 0.20*
G2: 0.17*
G2: 0.27*
updating
no. words OPINION
R^2 G2 = 0.019 (ns)

(b)

inhibition
G4: -0.20*
no. words NARR
G4: -0.15*
spelling
R^2 G4 = 0.166*
G4: -0.15*
G4: -0.29*
G4: 0.20*
G4: 0.44*
updating
no. words OPINION
R^2 G4 = 0.115*

(c)

inhibition
G8: -0.18*
no. words NARR
G8: -0.15*
G8: 0.19*
spelling
R^2 G8 = 0.304*
G8: -0.15*
G8: -0.37*
G8: 0.43*
G8: 0.23*
G8: 0.35*
updating
no. words OPINION
R^2 G8 = 0.261*

Table 5 Unstandardized coefficients and standard errors of direct and indirect effects on text generation at the word level

Effect	No. words, NA				No. words, OP			
	G2 Est (SE)	G4 Est (SE)	G8 Est (SE)	Significant differences	G2 Est (SE)	G4 Est (SE)	G8 Est (SE)	Significant differences
Direct effects								
HW	1.83 (0.99)	4.29* (1.03)	2.18* (1.09)	G2=G4=G8 / G2=G8	0.63 (0.45)	1.49* (0.46)	0.66 (0.69)	G2=G4=G8 / G2=G4
Spell	0.13 (0.29)	0.02 (0.45)	2.63* (0.87)	G2=G4<G8 / G2<G8	0.10 (0.14)	0.11 (0.21)	1.91* (0.49)	G2=G4<G8 / G2<G8
Inhibition M: HW	−0.27* (0.11)	−0.91* (0.29)	−2.19* (0.75)	G2<G4=G8 / G2<G8	−0.07 (0.06)	−0.34* (0.14)	−1.34* (0.41)	G2=G4<G8 / G2<G8
Inhibition M: spell	−0.35* (0.11)	−1.29* (0.32)	−2.26* (0.73)	G2<G4=G8 / G2<G8	−0.09 (0.06)	−0.46* (0.14)	−1.18* (0.38)	G2<G4=G8 / G2<G8
Updating M: HW	2.43* (0.77)	1.81 (1.13)	2.85* (1.82)	G2=G4=G8 / G2=G8	0.29 (0.30)	0.52 (0.56)	3.34* (1.16)	G2=G4<G8 / G2<G8
Updating M: spell	2.58* (0.77)	2.15 (1.19)	1.19 (1.95)	G2=G4=G8 / G2=G8	0.29 (0.31)	0.58 (0.57)	1.79 (1.27)	G2=G4=G8 / G2=G8
Indirect effects								
Inhib→HW	−0.06 (0.04)	−0.42* (0.13)	−0.53* (0.28)	G2<G4=G8 / G2=G8	−0.02 (0.02)	−0.15* (0.05)	−0.16 (0.17)	G2<G4=G8 / G2=G8
Inhib→spell	<0.01 (0.01)	<0.01 (0.06)	−0.47* (0.22)	G2=G4<G8 / G2<G8	<0.01 (0.01)	<0.01 (0.03)	−0.34* (0.15)	G2=G4<G8 / G2<G8
Upd→HW	0.22 (0.15)	0.32 (0.26)	0.82 (0.48)	G2=G4=G8 / G2=G8	0.07 (0.07)	0.11 (0.10)	0.25 (0.28)	G2=G4=G8 / G2=G8
Upd→spell	0.10 (0.24)	0.01 (0.22)	2.75* (1.05)	G2=G4<G8 / G2<G8	0.08 (0.11)	0.05 (0.11)	2.00* (0.62)	G2=G4<G8 / G2<G8

NA=narrative text; OP=opinion essay; Inhib→HW=indirect effect of inhibition via handwriting; Inhib→spell=indirect effect of inhibition via spelling; Upd→HW=indirect effect of updating via handwriting; Upd→spell=indirect effect of updating via spelling; M=mediator

$*p < .05$

Table 6 Unstandardized coefficients and standard errors of indirect effects on text generation at the word level. Partial mediation model

	No. words, NA				No. words, OP			
Inhib→HW	−0.09*	−0.27*	−0.67*	G2<G4<G8	−0.03*	−0.09*	−0.22*	G2<G4<G8
	(0.03)	(0.07)	(0.17)	G2<G8	(0.01)	(0.03)	(0.08)	G2<G8
Upd→HW	0.33*	0.21	1.04*	G2=G4<G8	0.11*	0.07	0.35*	G2<G4<G8
	(0.13)	(0.15)	(0.34)	G2<G8	(0.05)	(0.06)	(0.01)	G2<G8

NA=narrative text; OP=opinion essay; Inhib→HW=indirect effect of inhibition via handwriting;; Upd→HW=indirect effect of updating via handwriting

*p<.05

Importantly, while handwriting had a similar explanatory weight across grades, spelling was most relevant at G8.

Discussion

This paper sought to investigate the role of low-level executive functions and of transcription in text generation at three distinct levels of writing expertise: novice (G2), intermediate (G4), and upper-intermediate (G8) writers. The main goal of the study was to examine these relationships over a wide developmental window, so as to determine when such constraints predominate and how they interact with one another. Our first research question was related to the role of transcription in text generation. We had hypothesized that, with schooling (and writing) experience, children would progressively automatize transcription skills, to the point that they would explain little (if any) variance in text generation. Contrary to our prediction, and even though we did observe a significant improvement in both handwriting and spelling as a function of grade level, transcription skills significantly constrained text generation at all levels of writing or schooling experience. What is more, a comparison of the influence of transcription between grades 2, 4, and 8, revealed that the impact of transcription on text generation was the same at all grade levels. Other studies had also found that transcription constrained text generation at the word level, beyond the early primary grades (Abbott et al., 2010; Graham & Perin, 2007; Limpo & Alves, 2013a; Limpo, Alves, & Connelly, 2017; Wagner et al., 2011). However, a novel contribution of the present study is the insight that the weight of transcription in text generation is best explained as remaining the same throughout such an extended period of time. A few nuances need to be noted, however. First, handwriting had a more prominent role in text generation than spelling, in line with previous studies that had also found handwriting to explain more variance than spelling beyond the early primary grades (e.g., Graham et al., 1997). Nevertheless, spelling became relevant in the later grades at the word and sentence level. This could be due to the fact that spelling is a primary concern for teachers and students alike, and expectations about orthographic accuracy increase in the later years of primary and in secondary education; hence its heightened importance in the older groups of the study. Second, Catalan is a more consistent orthography than

Fig. 3 Standardized coefficients of the path analysis of the influence of EF measures on text generation at ▶ the sentence level at **a** Grade 2, **b** Grade 4, and **c** Grade 8, with handwriting as the mediator, and controlling for the effect of SES and sex (not shown, flor clarity). *NARR* narrative text, *OPINION* opinion-essay. *$p < .05$

English, so children may acquire a relatively high degree of mastery of spelling earlier than the English-speaking participants of other studies and, therefore, it might not constrain text generation as much in novice writers, which would explain why spelling did not account for variance before G8. Another reason for such a stable influence of transcription on text generation is that our G8 sample consisted of children who came from an overall lower socio-economic background, so their writing profile could be more similar to that of younger children (Hillocks, 2006). Finally, the constraints of transcription on text generation occurred mostly at the word level, in contrast to previous studies (e.g., Drijbooms et al., 2015), which found that it also constrained sentence-level text generation. However, Drijbooms et al. (2015) had only tested the effect of EFs on narrative writing in G4. This means that transcription constraints might not operate uniformly across genres or grade levels (Altemeier et al. 2006; Berninger et al., 2010).

Our second research question was whether executive functions influenced text generation at the word and sentence level directly or indirectly, via transcription. We had reasoned that, as long as transcription skills were not automatized, they would moderate at least part of the effect of EFs on text generation, but that, as they got mastered, EFs would support text generation directly. In this sense, and in line with previous studies (e.g., Drijbooms et al., 2015; Hooper et al., 2002), both inhibition and updating of WM had a direct and an indirect effect on text generation at the word level. The indirect effects of EFs through transcription are evidence of the fact that both spelling and, particularly, handwriting, limit the amount of text generated while they are themselves supported by these core low-level cognitive skills. This is in line with studies that found EFs to explain variance in levels of handwriting and spelling skills (e.g., Berninger et al., 2010). Inhibition may support handwriting by suppressing alternative letters or letter forms, or by inhibiting prepotent responses of competing motor programs. It may also support spelling by suppressing alternative spelling patterns, especially in cases of ambiguity (i.e., when more than one orthographic representation yields a phonologically plausible word form). Inhibition should be especially critical in a population of bilingual children like ours, who might often need to suppress correspondences in Spanish. Finally, updating of WM would serve to maintain a phonological form active in STM while the subject applies different orthographic rules or patterns (Berninger & Richards 2010), though this may be more typical of older participants; younger children are usually content with representing in writing the phonological structure of words (Salas, 2014).

Inhibition and updating should also impact text generation directly. Given that transcription constraints were operative at all grade levels, we would speculate that text generation for the children in our sample was accomplished using a knowledge-telling approach (Bereiter & Scardamalia, 1987). In contrast to more proficient writing (i.e., knowledge transforming), this type of text composition is not much
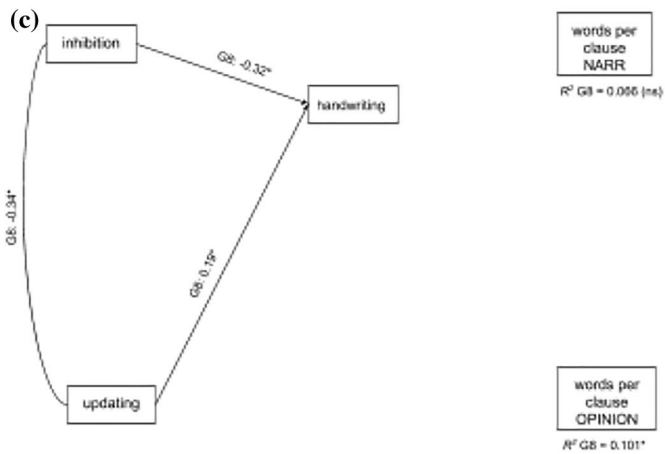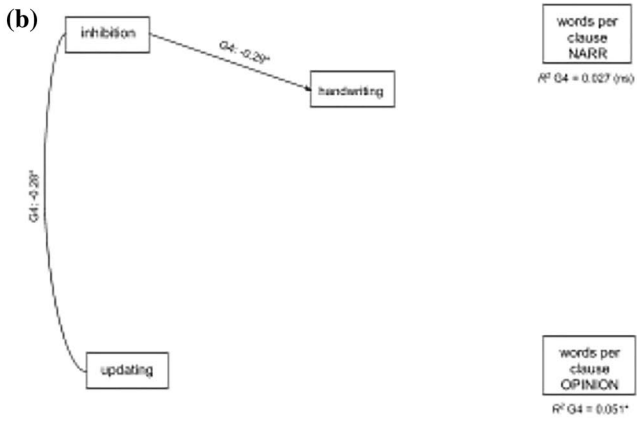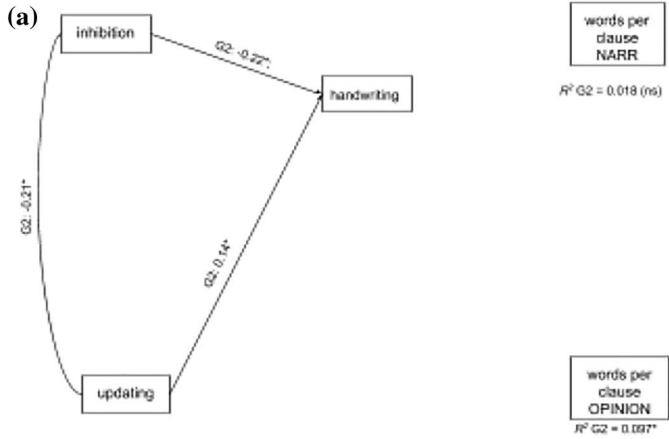
(a)

inhibition

G2: -0.22*

handwriting

G2: -0.21*

G2: 0.14*

updating

words per clause
NARR

R^2 G2 = 0.018 (ns)

words per clause
OPINION

R^2 G2 = 0.097*

(b)

inhibition

G4: -0.29*

handwriting

G4: -0.28*

updating

words per clause
NARR

R^2 G4 = 0.027 (ns)

words per clause
OPINION

R^2 G4 = 0.051*

(c)

inhibition

G8: -0.32*

handwriting

G8: -0.34*

G8: 0.15*

updating

words per clause
NARR

R^2 G8 = 0.066 (ns)

words per clause
OPINION

R^2 G8 = 0.101*

Table 7 Unstandardized coefficients and standard errors of direct and indirect effects on text generation at the sentence level

Effect	Words/clause, NA				Words/clause, OP			
	G2 Est (SE)	G4 Est (SE)	G8 Est (SE)	Significant differences	G2 Est (SE)	G4 Est (SE)	G8 Est (SE)	Significant differences
Direct effects								
HW	0.04 (0.03)	−0.03 (0.02)	0.01 (0.02)	G2=G4=G8 G2=G8	0.04 (0.04)	0.01 (0.02)	0.01 (0.02)	G2=G4=G8 G2=G8
Spell	0.02 (0.01)	0.01 (0.01)	0.01 (0.01)	G2=G4=G8 G2=G8	−0.01 (0.01)	<.01 (0.01)	0.01 (0.02)	G2=G4=G8 G2=G8
Inhibition M: HW	−0.01 (0.01)	0.01 (0.01)	<.01 (0.02)	G2=G4=G8 G2=G8	−0.01 (0.01)	<.01 (0.01)	−0.01 (0.01)	G2=G4=G8 G2=G8
Inhibition M: spell	−0.01 (0.01)	0.01 (0.01)	−0.01 (0.01)	G2=G4=G8 G2=G8	−0.01 (0.01)	<.01 (0.01)	−0.01 (0.01)	G2=G4=G8 G2=G8
Updating M: HW	−0.02 (0.04)	0.01 (0.02)	−0.01 (0.02)	G2=G4=G8 G2=G8	−0.05 (0.04)	−0.01 (0.03)	0.01 (0.03)	G2=G4=G8 G2=G8
Updating M: spell	−0.03 (0.04)	<.01 (0.02)	<0.01 (0.03)	G2=G4=G8 G2=G8	−0.05 (0.04)	−0.01 (0.03)	0.01 (0.04)	G2=G4=G8 G2=G8
Indirect effects								
Inhib→HW	−0.01 (0.01)	0.01 (0.01)	−0.02* (0.01)	G2=G4<G8	−0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)	G2=G4=G8
Inhib→spell	0.00 (0.00)	−0.01* (0.01)	−0.01 (0.01)	G2<G4=G8	0.00 (0.00)	0.00 (0.00)	−0.01 (0.01)	G2=G4=G8
Upd→HW	0.01 (0.01)	−0.01 (0.01)	0.04* (0.02)	G2=G4<G8	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	G2=G4=G8
Upd→spell	0.02 (0.01)	0.02* (0.01)	0.04 (0.03)	G2=G4=G8	−0.01 (0.01)	−0.01 (0.01)	0.03 (0.02)	G2=G4=G8

NA=narrative text; OP=opinion essay; Inhib→HW=indirect effect of inhibition via handwriting; Inhib→spell=indirect effect of inhibition via spelling; Upd→HW=indirect effect of updating via handwriting; Upd→spell=indirect effect of updating via spelling; M=mediator

*$p < .05$

concerned with rhetorical issues. Accounts of types of knowledge-telling processes (Berninger, Fuller, & Whitaker 1996; Hayes, 2011) describe a writer who needs to keep the topic of the text active in mind, which acts as a cue to search for content in LTM. Once content is retrieved, simple adequacy tests ensue (e.g., *Does this idea fit the topic?*). Arguably, keeping the topic active whilst searching for content requires updating of WM, while performing adequacy tests requires inhibition skills to suppress content that is not a good fit for the text.

In contrast to the word level, EFs had very little impact on text generation at the sentence level. We have already explained some of the discrepancies with previous studies (e.g., Drijbooms et al., 2015). In addition, our measure of sentence-level text generation, while comparable and equally popular (e.g., Tolchinsky & Salas, 2018), was slightly different: we measured number of words per clause, whereas Drijbooms et al. (2015) measured number of words per T-unit, where a T-unit is the main clause plus all of its subordinate clauses (Hunt, 1965). We chose against identifying T-units because they do not provide a clear rationale for certain clauses, particularly juxtaposed clauses, and because they may not be adequate to accommodate some characteristics typical of discourse development[5] (Aparici, 2010). Our results are thus more in line with previous research on the development of syntactic complexity in words per clause, which has indicated that there is little variation over time and that the increase in text length as a function of grade level is not necessarily accompanied by an increase in clause length (Llauradó-Singla, 2012). Moreover, the low developmental and inter-individual variation on this measure (the mean at all grade levels ranged only between 4 and 5.5 words per clause, and standard deviations were as low as between ± 1.26 and ± 1.80 words) might explain why we found virtually no significant effects on text generation at this level.

Our study supports the view that updating of WM and inhibition are involved in written composition, especially at the word level, both directly and indirectly. Considering that these two skills were significantly correlated across grade levels and models, while the pattern of relationships with other variables was remarkably similar as well, they appear to be jointly providing support to both low- and high-level writing skills.

Limitations

The present study could be improved in several ways. First, a more complete picture would emerge by including high-level EF measures of planning and revising, so as to better understand the extent to which low-level EFs support them, and whether and how these relate to transcription, on the one hand, and to text generation, on the other. Similarly, because some high-level EF skills may only truly emerge as

---

[5] Some uses of *and* in children's spoken and written discourse may not equate to adults' use as a conjunction; thus, coordination in children could be easily overestimated. Conversely, subordination may be underestimated, since several children use *and* as well as juxtaposition where an adult would use a subordinating conjunction (see Aparici, 2010 for a detailed overview).

Fig. 4 Standardized coefficients of the path analysis of the influence of EF measures on text generation ▶
at the sentence level at **a** Grade 2, **b** Grade 4, and **c** Grade 8, with spelling as the mediator, and control-
ling for the effect of SES and sex (not shown, flor clarity). *NARR* narrative text, *OPINION* opinion-essay.
$*p < .05$

a consequence of explicit instruction (e.g., SRSD, Graham et al., 2012), it may be
useful to test these models before and after an intervention. Second, previous stud-
ies suggested that cognitive flexibility may not be entirely discernable from either
inhibition or working memory in children (e.g., Drijbooms et al., 2015; van der
Ven et al., 2013). However, it would be interesting to include cognitive flexibility
in future investigations that examine EFs in samples including a wider age range
than the present study. Third, having established that most (or all) of our partici-
pants still generate text following a knowledge-telling approach, a comparison with
more experienced or even expert writers should shed light on how EF (both low- and
high-level) support text generation. Finally, we have only examined basic features
of text generation, but future research should strive to understand how specific text
features (e.g., coherence, vocabulary precision and richness, among others) relate to
and are supported by low-level EFs.

Concluding remarks

We have presented compelling evidence that transcription skills, particularly hand-
writing, constrain text generation at least up to the early secondary education years
in a semi-consistent orthography. An important educational implication of this find-
ing is that handwriting fluency should be attended to even in the later elementary
school years. Also, while spelling should remain a focus of teachers throughout
elementary education, educators should carefully consider its impact on children's
developing writing processes. This means that improving spelling accuracy should
not be at the expense of making students focus mostly on spelling, thus prevent-
ing them from devoting attention to other writing processes and features. Teachers
should thus provide students with strategies to cope with the complexity of writing;
for example, guiding students to focus on spelling only at a specific revision phase,
thus lessening cognitive demands for the rest of the composition process. We have
also provided evidence that two of the core EFs, inhibition and updating of WM,
support both transcription skills and text generation at the word level over a long
developmental period. The fact that these have also been found in other languages
(e.g., English, Dutch) points to them as skills that may underlie written compo-
sition universally, thus contributing to a line of research that claims that literacy
learning across languages involves the recruitment of the same cognitive mecha-
nisms, despite marked differences in developmental rates (Caravolas et al., 2012;
2013).
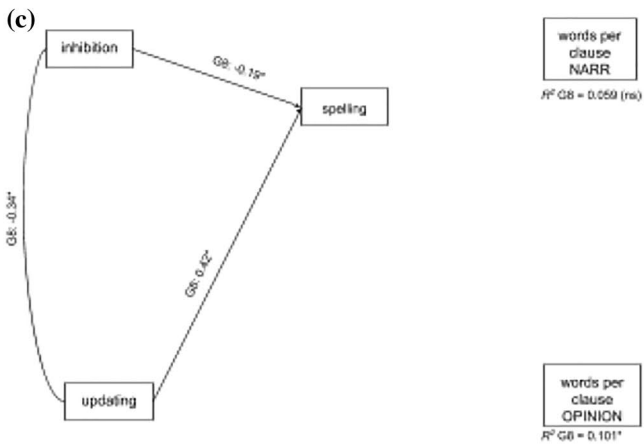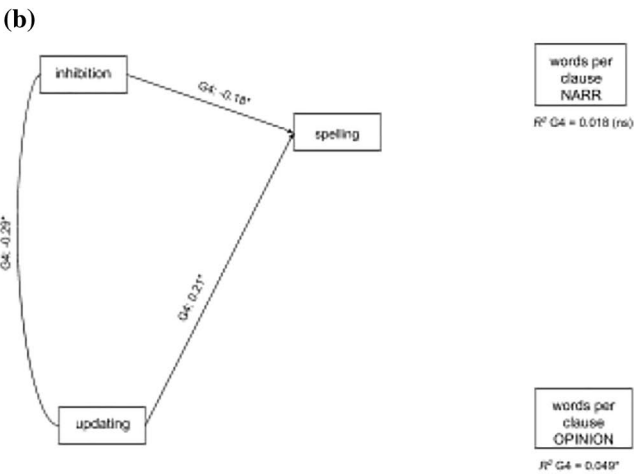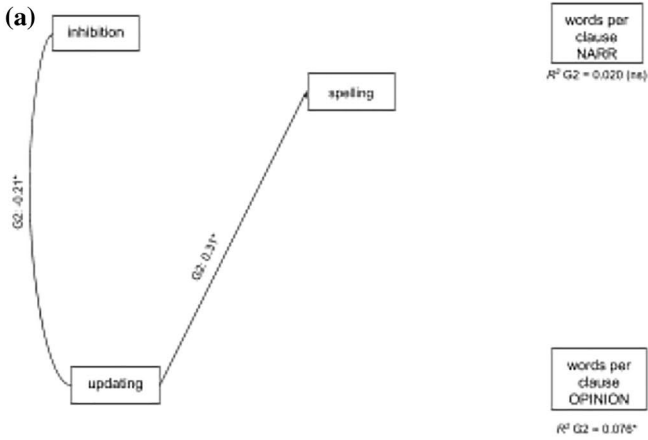
(a)

(b)

(c)

References

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281–298.

Ailhaud, E., Chenu, F., & Jisa, H. (2016). A developmental perspective on the units of written French. In J. Perera, M. Aparici, E. Rosado, & N. Salas (Eds.), *Written and spoken language development across the lifespan* (pp. 287–305). Cham: Springer.

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.

Altemeier, L., Jones, J., Abbott, R. D., & Berninger, V. W. (2006). Executive functions in becoming writing readers and reading writers: Note taking and report writing in third and fifth graders. *Developmental Neuropsychology, 29*(1), 161–173.

Alves, R. A., & Limpo, T. (2015). Progress in written language bursts, pauses, transcription, and written composition across schooling. *Scientific Studies of Reading, 19*(5), 374–391.

Aparici, M. (2010). *El desarrollo de la conectividad discursiva en diferentes géneros y modalidades de producción.* Unpublished doctoral dissertation. Universitat de Barcelona.

Babayigit, S., & Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: A longitudinal study. *Reading and Writing: An Interdisciplinary Journal, 23*(5), 539–568.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

Berman, R., & Nir-Sagiv, B. (2009). Cognitive and linguistic factors in evaluating text quality: Global versus local? In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics* (pp. 421–440). Amsterdam: John Benjamins.

Berman, R., & Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: Speech and writing. *Written Language & Literacy, 5*(1), 1–43.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A cross linguistic developmental study*. Hillsdale, NJ: Laurence Erlbaum.

Berninger, V., Abbott, R., Cook, C. R., & Nagy, W. (2017). Relationships of attention and executive functions to oral language, reading, and writing skills and systems in middle childhood and early adolescence. *Journal of Learning Disabilities, 50*(4), 434–449.

Berninger, V., Whitaker, D., Feng, Y., Swanson, H. L., & Abbott, R. D. (1996). Assessment of planning, translating, and revising in junior high writers. *Journal of School Psychology, 34*(1), 23–52.

Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing, 4*(3), 257–280.

Berninger, V. W., Abbott, R. D., Swanson, H. L., Lovitt, D., Trivedi, P., Lin, S. J. C., et al. (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., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 345–363). New York, NY: Guilford Press.

Berninger, V. W., Cartwright, A. C., Yates, C. M., Swanson, H. L., & Abbott, R. D. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades. *Reading and Writing, 6*(2), 161–196.

Berninger, V. W., Fuller, F., & Whitaker, D. (1996). A process model of writing development across the life span. *Educational Psychology Review, 8*(3), 193–218.

Berninger, V., & Richards, T. (2010). Inter-relationships among behavioral markers, genes, brain and treatment in dyslexia and dysgraphia. *Future Neurology, 5*(4), 597–617.

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

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review, 29*(3), 180–200.

Bryson, M., Bereiter, C., Scardamalia, M., & Joram, E. (1991). Going beyond the problem as given: Problem solving in expert and novice writers. In R. Stumberg & P. Frensch (Eds.), *Complex problem solving*. Hillsdale, NJ: Lawrence Erlbaum.

Caravolas, M., & Bruck, M. (1993). The effect of oral and written language input on children's phonological awareness: A cross-linguistic study. *Journal of Experimental Child Psychology, 55*, 1–1.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*(8), 1398–1407.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., et al. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science, 23*(6), 678–686.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255.

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.

Cordeiro, C., Castro, S. L., & Limpo, T. (2018). Examining potential sources of gender differences in writing: The role of handwriting fluency and self-efficacy beliefs. *Written Communication, 35*(4), 448–473.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64,* 135–168.

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.

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socioeconomic index of occupational status. *Social Science Research, 21*(1), 1–56.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology, 89*(1), 170–182.

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476.

Hajovsky, D. B., Villeneuve, E. F., Reynolds, M. R., Niileksela, C. R., Mason, B. A., & Shudak, N. J. (2018). Cognitive ability influences on written expression: evidence for developmental and sex-based differences in school-age children. *Journal of School Psychology*, *67*, 104–118.

Hayes, J. R. (2011). Kinds of knowledge-telling: Modeling early writing development. *Journal of Writing Research, 3*(2), 73–92.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369–388.

Hillocks, G. (2006). Middle and high school composition. In P. Smagorinsky (Ed.) *Research on composition: Multiple perspectives on two decades of change* (pp. 48–77).

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In I.W. Gregg & E.R. Steinberg (Eds.). Cognitive processes in writing. Hillsdale. N.J.: Erlbaum, pp. 3–30.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16*(3), 174–180.

Hooper, S. R., Swartz, C. W., Wakely, M. B., De Kruif, R. E., & Montgomery, J. W. (2002). Executive functions in elementary school children with and without problems in written expression. *Journal of Learning Disabilities, 35*(1), 57–68.

Hoskyn, M., & Swanson, H. L. (2003). The relationship between working memory and writing in younger and older adults. *Reading and Writing, 16*(8), 759–784.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (No. 3). Champaign, IL: National Council of Teachers of English.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78,* 243–255.

Kellogg, R. T. (1994). *The psychology of writing*. New York: Oxford University Press.

Kellogg, R. T. (2001). Competition for working memory among writing processes. *American Journal of Psychology, 114*(2), 175–192.

Kent, S., Wanzek, J., Petscher, Y., Al Otaiba, S., & Kim, Y.-S. (2014). Writing fluency and quality in kindergarten and first grade: The role of attention, reading, transcription, and oral language. *Reading and Writing, 27*(7), 1163–1188.

Kim, Y.-S., Al Otaiba, S., Folsom, J. S., & Greulich, L. (2013). Language, literacy, attentional behaviors, and instructional quality predictors of written composition for first graders. *Early Childhood Research Quarterly, 28*(3), 461–469.

Kim, Y.-S., Al Otaiba, S., Folsom, J. S., Greulich, L., & Puranik, C. (2014). Evaluating the dimensionality of first-grade written composition. *Journal of Speech, Language, and Hearing Research, 57*(1), 199–211.

Kim, Y.-S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: An exploratory study. *Learning and Individual Differences, 21*(5), 517–525.

Kline, R. B. (2011). *Methodology in the Social Sciences. Principles and practice of structural equation modeling*. New York: Guilford Press.

Lerkkanen, M.-K., Rasku-Puttonen, H., Aunola, K., & Nurmi, J.-E. (2004). The developmental dynamics of literacy skills during the first grade. *Educational Psychology, 24*(6), 793–810.

Limpo, T., & Alves, R. A. (2013a). Modeling writing development: Contribution of transcription and self-regulation to Portuguese students' text generation quality. *Journal of Educational Psychology, 105*(2), 401–413.

Limpo, T., & Alves, R. A. (2013b). Teaching planning or sentence-combining strategies: Effective SRSD interventions at different levels of written composition. *Contemporary Educational Psychology, 38*(4), 328–341.

Limpo, T., Alves, R. A., & Connelly, V. (2017). Examining the transcription-writing link: Effects of handwriting fluency and spelling accuracy on writing performance via planning and translating in middle grades. *Learning and Individual Differences, 53*, 26–36.

Llauradó-Singla, A. (2012). *The relation between lexicon and syntax in texts written in Catalan by school children and adolescents*. Unpublished doctoral dissertation. Universitat de Barcelona.

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist, 35*(1), 13–23.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568.

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.

Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30*(3), 308–330.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus User's Guide. Eighth Edition.* Los Angeles, CA: Muthén & Muthén.

Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high-and low-level production processes in written composition. *Memory & Cognition*, *30*(4), 594–600.

Olive, T., Favart, M., Beauvais, C., & Beauvais, L. (2009). Children's cognitive effort and fluency in writing: effects of genre and of handwriting automatisation. *Learning and Instruction*, *19*(4), 299–308.

Puranik, C. S. (2006). Expository writing skills in elementary school children from third through sixth grades and contributions of short-term and working memory. Unpublished doctoral dissertation. University of Florida.

Puranik, C. S., & Al Otaiba, S. (2012). Examining the contribution of handwriting and spelling to written expression in kindergarten children. *Reading and Writing, 25*(7), 1523–1546.

Puranik, C. S., Lombardino, L. J., & Altmann, L. J. P. (2008). Assessing the microstructure of written language using a retelling paradigm. *American Journal of Speech-Language Pathology, 17*(2), 107–120.

Salas, N. (2014). *Early development of text writing in two contrasting Orthographies: English and Spanish*. Unpublished doctoral dissertation. School of Psychology, Bangor University.

Salas, N. & Caravolas, N. (submitted). Commonalities in early writing outcomes and predictors across English and Spanish.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods, 8*(2), 597–599.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.

Swanson, H. L., & Berninger, V. W. (1996). Individual differences in children's working memory and writing skill. *Journal of Experimental Child Psychology, 63*(2), 358–385.

Tolchinsky, L., & Salas, N. (2018). Grammar for writing and grammar in writing. In A. Bar-On & D. Ravid (Eds.), *Handbook of communication disorders. Theoretical, empirical, and applied linguistics perspectives*. Berlin: de Gruyter.

Vanderberg, R., & Swanson, H. L. (2007). Which components of working memory are important in the writing process? *Reading and Writing*, *20*(7), 721–752.

van der Ven, S. H., Kroesbergen, E. H., Boom, J., & Leseman, P. P. (2013). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *British Journal of Developmental Psychology, 31*(1), 70–87.

Vandenbroucke, L., Verschueren, K., & Baeyens, D. (2017). The development of executive functioning across the transition to first grade and its predictive value for academic achievement. *Learning and Instruction, 49,* 103–112.

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., et al. (2011). Modeling the development of written language. *Reading and Writing, 24*(2), 203–220.

Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. Chichester: Wiley.

Wechsler, D. (2005). *Wechsler intelligence scale for children-fourth edition, Spanish*. San Antonio, TX: Harcourt Assessment.