the british psychological society
promoting excellence in psychology

British Journal of Educational Psychology (2021), 91, 714–734

www.wileyonlinelibrary.com

Child-level factors affecting rate of learning to write in first grade

Mark Torrance[1]*, María Arrimada[2] and Sarah Gardner[1]

[1]Department of Psychology, School of Social Sciences, Nottingham Trent University, Nottingham, UK

[2]Departmento de Psicología, Sociología y Filosofía, Facultad de Educación, Universidad de León, León, Spain

Abstract. **Background.** Written composition requires handwriting, spelling, and text planning skills, all largely learned through school instruction. Students' rate of learning to compose text in their first months at school will depend, in part, on their literacy-related abilities at school start. These effects have not previously been explored.

Aim. We aimed to establish the effects of various literacy-related abilities on the learning trajectory of first-grade students as they are taught to write.

Sample. 179 Spanish first-grade students (94 female, mean age 6.1 years) writing 3,512 texts.

Method. Students were assessed at start of school for spelling, transcription fluency, letter knowledge, phonological awareness, handwriting accuracy, word reading, and non-verbal reasoning. They were then taught under a curriculum that included researcher-designed instruction in handwriting, spelling, and ideation. Students' composition performance was probed at very regular intervals over their first 13 weeks at school.

Results. Controlling for age, overall performance was predicted by spelling, transcription fluency, handwriting accuracy, word reading, and non-verbal reasoning. Most students showed rapid initial improvement, but then much slower learning. Weak spellers (and to a lesser extent less fluent hand-writers) showed weaker initial performance, but then steady improvement across the study period.

Conclusion. Transcription ability at school entry affects response to writing instruction.

Most children start school with a well-developed ability to compose their thoughts in speech, but without the ability to compose their thoughts in writing. Most obviously, this is because the ability to spell and handwrite requires explicit instruction, and in most educational systems, this instruction does not start in earnest until the beginning of first grade. Written composition also requires a different approach to retrieving ideas. A parent or teacher asking a child to speak a story can interject with 'Where were they?', 'What

*Correspondence should be addressed to Mark Torrance, Psychology Department, School of Social Sciences, Nottingham Trent University, Shakespeare Street, Nottingham NG1 4FQ, UK (email: mark.torrance@ntu.ac.uk).

DOI:10.1111/bjep.12390

happened next?', and so forth. Writing a story requires that these cues are internalized: The child needs to take control of their own narrative-production strategy so that it can be generated without external prompts.

To compose text, therefore, students need to master both transcription (spelling and handwriting) and ideation (generating and structuring relevant content) – the 'simple view of writing' (Juel, 1988; Juel, Griffith, & Gough, 1986) – and transcription and ideation need to be taught. Written production is, however, cognitively challenging: Simultaneous focus on ideation and transcription imposes considerable attentional demands (McCutchen, 1996; Torrance & Galbraith, 2006), and this is particularly the case when both ideation and transcription skills are not well developed. The 'not so simple view of writing' is that, for these reasons, learning to write also requires that students acquire self-regulatory strategies for controlling how they allocate attention (Berninger & Winn, 2006; see also Kim & Park, 2019).

Students starting school therefore face a formidable set of challenges as they set out to learn how to compose text. Transcription will be far from automatized and so will draw attention away from ideation, and students will lack the necessary regulatory strategies to redress this balance. This will not only affect the quality of their text but also their ability to learn: the 'double challenge' of learning to write (Rijlaarsdam & Couzijn, 2000).

Our focus in this paper is students' rate of learning to compose text across their first semester at primary (elementary) school, and particularly how this is affected by literacy-related skills (spelling, transcription fluency, letter knowledge, phonological awareness, handwriting accuracy, word reading), measured at school entry, that students bring to their learning. Students' initial ability in each of these will depend on a range of factors, including parents' education, family size, and home literacy activity (Blatchford, 1991; Coker, 2006; Dunsmuir & Blatchford, 2004; Taylor & Schatschneider, 2010; Van Steensel, 2006). Students entering primary schools are possibly more developmentally heterogeneous than at any subsequent time in their school career.

Several basic abilities have been found to correlate with handwriting and spelling competence in children in kindergarten and the first three grades of primary school. Handwriting fluency and/or accuracy are predicted by visual-motor integration (Cornhill & Case-Smith, 1996; Daly, Kelley, & Krauss, 1994), motor coordination (Cornhill & Case-Smith, 1996; Frolek & Luze, 2014; Tseng & Murray, 1994), visual-motor integration (Tseng & Murray, 1994), and single-letter writing, vocabulary, and grammar (Kent, Wanzek, Petscher, Al Otaiba, & Kim, 2014). Spelling ability is predicted by phonological awareness (Babayiğit & Stainthorp, 2011; Frost, 2001; Lehtonen & Bryant, 2004; Mäki, Voeten, Vauras, & Poskiparta, 2001; Nation & Hulme, 1997), single-letter writing from a dictated letter name (Puranik, Lonigan, & Kim, 2011), knowledge of phoneme-grapheme correspondence (Sadoski, Willson, Holcomb, & Boulwar-Gooden, 2004), and short-term memory (Binamé & Poncelet, 2016). Production fluency measured as rate of handwritten alphabet recall and/or text copying predicts composition quality in second and third grade (Jones & Christensen, 1999; Kim, Al Otaiba, Wanzek, & Gatlin, 2015) after control for reading ability and several other literacy-related variables. A combination of spelling and reading ability predicts composition productivity in 5-year-old children (US kindergarten), measured as counts of words, sentences, and ideas (Kent et al., 2014).

However, fewer studies have explored predictors of composition quality[1] in early grades. A meta-analysis by Kent and Wanzek (2016) analysed effects of spelling ability and handwriting fluency on composition quality in children from kindergarten to third grade and found a significant mean effects in both cases (spelling, 6 studies, mean effect = .49; handwriting fluency, 7 studies, mean effect = .59). Evidence for the effects of spelling ability on composition quality specifically in first grade is limited and mixed. Jiménez and Hernández-Cabrera (2019) found that a composite spelling and handwriting factor predicted composition quality in Spanish first graders. In a much earlier study, Juel (1988) explored spelling effects on composition quality controlling for the quality of students' spoken narratives, as a writing-independent measure of content planning and structuring skills (i.e., the ideation component of the Simple View of Writing). Spelling ability was much stronger than ideation as a predictor of composition quality in first grade. Kim, Al Otaiba, Folsom, and Gruelich (2013), Kim, Otaiba, Folsom, Greulich, and Puranike (2014) in analyses that included measures of handwritten alphabet recall speed, passage comprehension, and spelling-to-dictation, found a clear effect for passage comprehension and a weak effect of alphabet writing, but the effect for spelling was not statistically significant. Jones and Christensen (1999, Study 1) found a strong relationship between handwritten alphabet recall speed and composition quality, controlling for reading ability, but not spelling. Wagner et al. (2011) found that handwriting fluency (a combined sentence-copying and alphabet recall factor) predicted quality of students' macro-structure and syntactic complexity, but spelling and punctuation (aggregated) in an expository writing task, with no control for spelling or other factors, did not.

Several other factors less directly involved in written production predict early composition quality. Kent and Wanzek (2016) identified 12 studies that examined correlation with various measures of reading ability in children in the first 3 years of school and found a mean effect of 0.48. In just first grade, Abbott and Berninger (1993) found that a composite measure of mainly word and non-word naming predicted composition quality (in the absence of control for spelling). Other studies have found correlations with ability to maintain attention (Kent et al., 2014; Kim et al., 2013, 2015) and verbal and non-verbal general ability measures (Olinghouse, 2008).

All of the findings cited thus far are correlations between various factors and writing quality measured at a single point in time. Their focus therefore is on predicting composition performance rather than composition learning. A small handful of studies has examined the effects of spelling and handwriting ability at or prior to school entry on children's composition performance at a later time-point. Dunsmuir and Blatchford (2004) found that handwriting accuracy at school entry predicted writing performance at the end of second grade. Kent et al. (2014) found that spelling and reading measures in kindergarten predict composition quality in first grade even after control for kindergarten written productivity. However, they did not find a relationship between handwritten alphabet recall in kindergarten and first-grade composition quality. Mäki et al. (2001) found that spelling ability in first grade predicted coherence of text written in second grade. Babayiğit and Stainthorp (2010) found that spelling-to-dictation in first grade predicted content and structure of compositions written a year later. It is worth noting that these four studies, conducted in the UK, USA, Finland, and Cyprus, respectively,

[1] *We use the terms 'composition quality' and 'composition performance' to refer generally to all features of a text – typically a narrative in the studies that we cite – that makes it coherent and meaningful, including both ability to handwrite, spell and construct sentences, and ability to generate and structure content.*

represent four different educational systems and languages with both deep and shallow orthographies.

Finding effects of handwriting and spelling on later performance suggests that that these factors might have an effect on how well students respond to writing instruction. The study presented in this paper aimed to provide a more robust test of this hypothesis. Specifically, we explored the effects of literacy-related factors measured at school entry on the timecourse of writing development, inferred from composition performance measured at a large number of time-points over the subsequent semester. This provided both a more robust overall measure of students' writing ability – single-point measures of writing performance are notoriously noisy (Van den Bergh, Maeyer, van Weijen, & Tillema, 2012) and, more importantly, permitted estimation of changes in students rate of learning over time.

Our aim therefore was to determine how the literacy-related abilities that a child brings with them into first grade affects the rate at which their written composition ability then develops. At school entry (start of 1st grade), students completed a battery of literacy-related tests. These included measures of handwriting accuracy and fluency and of spelling accuracy. Then over the following 13 weeks, they completed regular (at least weekly) narrative writing tasks. During this time, students were taught according to a curriculum that included researcher-prescribed instruction in text planning (idea generation and organization) and both researcher-prescribed and normal-curriculum instruction in spelling and handwriting. We anticipated, and found, considerable variation across students both in overall learning and in learning rate. Our study determined whether, after control for age, non-verbal ability, and various literacy-related measures not directly associated with transcription, spelling and handwriting ability at school entry affected the subsequent timecourse of students' learning to compose text.

Method

Participants

Participants comprised all students in eight first-grade classes, each with different classroom teachers, distributed across three *concertados* schools in middle-class areas of León (Spain). Two students were dropped from the sample because they showed substantial developmental delay and/or very poor attendance, giving $N = 179$ (94 female). Mean age at the beginning of first grade was 6.1 years ($SD = 3.37$). All participants spoke Spanish as their home language.

Educational and instructional context

In the Spanish educational system, students start primary (elementary) school in the year that they reach 6 years of age. Younger children have the option of attending kindergarten, and this was the case for all students in our sample. Writing instruction in kindergarten focuses exclusively on transcription, with no reference to composition or text quality. At the end of kindergarten, most students are able to name, sound, and form all letters. It is also expected that students leave kindergarten with knowledge of phoneme-grapheme correspondence and able to write syllables and simple words. Some students, though not the majority, are able to write simple short sentences.

All students in our sample were taught in single-teacher classroom groups. Teachers (two male) had teaching experience, within the Spanish school system, ranging between 11 and 39 years.

All students in our sample took part in a programme of researcher-designed written composition instruction that aimed to teach both transcription and ideation, based on methods previously evaluated as successful, and described in detail in Arrimada, Torrance, and Fidalgo (2018a, 2018b). This provided some standardization of instruction against which to interpret students' learning. Students received three 15-min sessions per week in which they completed exercises in one of spelling, handwriting, sentence-combining, and on strategies for developing the content and structure of their narratives. These tasks commenced in the first week of the 13-week period in which we assessed students' composition performance and continued throughout.

Instructional tasks were introduced and overseen by classroom teachers, all of whom received an initial training session, and also had brief, weekly trouble-shooting meetings with the lead researcher. Alongside, these sessions' teachers continued with their normal classroom curricula. This focused almost exclusively on handwriting and spelling with children writing words and simple sentences, mostly to dictation and by copying.

Measures

In the first 3 weeks of school, we delivered a battery of tests assessing a range of skills. We then assessed written composition performance at multiple time-points over the following 13 weeks.

Start-of-year measures

Tasks were administered in whole-class groups by the lead researcher across three sessions lasting between 20 and 50 min.

Spelling. Students completed real and pseudoword spelling-to-dictation tasks. *Real Words:* We selected 12 bisyllabic and trisyllabic medium-frequency words from the Spanish dictionary of word frequency in children's writing (Martínez & García, 2004). Following Defior, Jiménez-Fernández, and Serrano (2009), each word included some form of spelling difficulty. *Pseudowords* were matched to real words in syllabic and phonemic structure. Words lists and more detailed explanation are provided in Appendix A. Words were analysed for the number of errors, any of substitution with the wrong letter, omission, or position swapping counting as a single error. Responses scored 2 points if correct, 1 point if one or two errors, and 0 if more than two errors. Both real and pseudowords were scored in the same way, with any phonologically plausible letter being counted as correct for pseudowords (but not for real words). Cronbach's alpha (from data collected in this study): Real words, 0.93, 95% CI [0.91, 0.94]; pseudowords, 0.92, 95% CI [0.90, 0.94].

Transcription fluency. Students completed two sentence-copying tasks (following, for example, Barnett, Henderson, Scheib, & Schulz, 2009) and copied the alphabet (e.g., Berninger et al., 1992). For the sentence tasks, students wrote a regularly spelled, easy to remember sentence (*Me gusta mucho salir al patio*/I really like going to the playground)

as many times as possible in 1 min. Students wrote in their neatest handwriting (*copy accurate*) and then as quickly as possible (*copy speed*). Score in both cases was the number of words written that was either correct or phonologically close to the target word. For the *alphabet-copy* task, students were given a copy of the alphabet and asked to handwrite it as many times as possible in 1 min, scoring one point for each recognizable letter.

Handwriting accuracy. Two raters scored handwriting accuracy – the extent to which letters were correctly and neatly shaped – for the first 10 words, or for the full text if less than 10 words, of a written narrative task (not one of the composition performance tasks detailed below). To reduce potential carry-over effects from other features of the text (spelling, content), one rater did not speak Spanish. Handwriting accuracy was scored from 0 (most marks on the page could not be identified as letter) to 4 (nearly all characters accurate and regular), details in Appendix B. Inter-rater agreement (intraclass correlation) was 0.86 (95% CI [0.78, 0.91]).

Letter knowledge. The researcher spoke the name of each letter of the alphabet twice, following the alphabet sequence. Students wrote down the letters they heard. Students scored a point for each identifiable, correct letter. Students were free to write in upper or lower case. Cronbach's alpha was 0.93 (95% CI [0.91, 0.94]).

Phonological awareness. Students were provided with sets of 30 pictures found in piloting to be easy to name. The researcher spoke a phoneme and gave two examples of words starting with that phoneme. Students were then given 30s to find as many pictures as possible with a name that started with the phoneme. This was repeated for 5 sets of pictures and using the phonemes /θ/ /k/ /g/ /f/ and /t/. Score was the total number of pictures identified correctly totalled across the 5 sets, with a maximum of 75 (15 per set). Internal reliability across the five phonemes/picture sets (Cronbach's alpha) was 0.88 (95% CI [0.86, 0.91]).

Word reading. Students were presented with three sheets giving 30 words each. Half the words in each sheet were nouns representing an object category (objects found on a farm, in a bedroom, or on the beach, with a different category on each sheet). Students were asked to circle as many words in this category as they could within 30 s. Students then repeated the same task, but this time with sheets showing pictures instead of words and with location on the page rearranged. Cronbach's alpha: word reading, 0.84 (95% CI [0.80, 0.88]); picture selection, .84 (95% CI [0.84, 0.88]). Correlation in scores between the word and corresponding picture versions of the task was 0.34, 0.32, and 0.22 for farm, bedroom, and beach. We then regressed total scores from the reading task onto total scores from the picture task. Residuals provided a direct measure of a child's fluency in single-word reading for meaning, controlling for students' domain knowledge and other non-reading task-specific abilities (general speed of processing, ability to sustain attention etc.).

Non-verbal reasoning. Students completed a matrix task designed as a shortened, group-administered version of Raven's progressive matrices (Raven, 1981). The task comprised 20 matrices split in three sets: patterns, sequences of identical figures, and geometrical figures. For each matrix, students circled the picture that completed the matrix among six options. Cronbach's alpha was 0.71, (95% CI [0.66, 0.77]).

Written composition

Composition performance – our dependent variable – was determined as follows: Students wrote a narrative describing events in their own lives (e.g., 'What I did yesterday', 'How I celebrated my last birthday') a task similar to that used by, for example, Kent et al. (2014). Tasks were administered by classroom teachers and had a time limit of 15 min. The first task was completed in the week following the end of initial testing – about 4 weeks after the start of school – and the final task was completed 13 weeks later. Students completed this narrative composition task at minimum once per week and in most cases twice per week. There was, however, some variation in the number of tasks completed across students, due to absence, and across classrooms.

Our sample comprised a total of 3,512 texts with a median of 22 texts per student (minimum = 9, maximum = 25). These were given a single holistic quality rating on a six-point scale, detailed in Appendix C. All texts were scored by the lead researcher. A second, trained rater scored a random sample of 19% (660 texts). Inter-rater reliability (intraclass correlation) was 0.92 (95% CI [0.90, 0.93]).

Results

As can be seen from Figure 1, the general trend across all students was for an initial period of rapid improvement in composition quality followed by a longer period of much more gradual improvement. This suggested that growth in performance was best modelled as two separate growth curves in a piecewise growth curve model (e.g., Chou, Yang, Pentz, & Hser, 2004). We first evaluated a series of models to establish which hinge-point – the test occasion that marked the boundary between the first and second timepieces – best fitted the data. We then compared the best fit model from this analysis with a model that hypothesized a single, linear growth trend. This analysis is described in the *Learning Rates* section that follows. We then evaluated a series of models to establish the effects of the various measures of literacy skill at start of school on overall performance across the study (main effects), on initial rate of learning (the first timepiece) and rate of later learning (the second timepiece). These analyses are reported in the *Factors affecting learning rates* section.

Modelling was by linear mixed-effects regression. Our data comprised clusters of observations for each child, and clusters of children nested within classroom. All models therefore estimated random intercepts for each classroom and for each child. They also estimated random slopes for each timepiece (initial and later growth curves) for classrooms and for children nested within classrooms. Table 2 makes this random effects structure clear. Models were evaluated with maximum likelihood estimation using the R lme4 package (Bates, Mächler, Bolker, & Walker, 2015). We established relative model fit by direct comparison of AIC for non-nested models and by likelihood-ratio chi-squared tests for nested models. The random effect structure just that we describe above gave

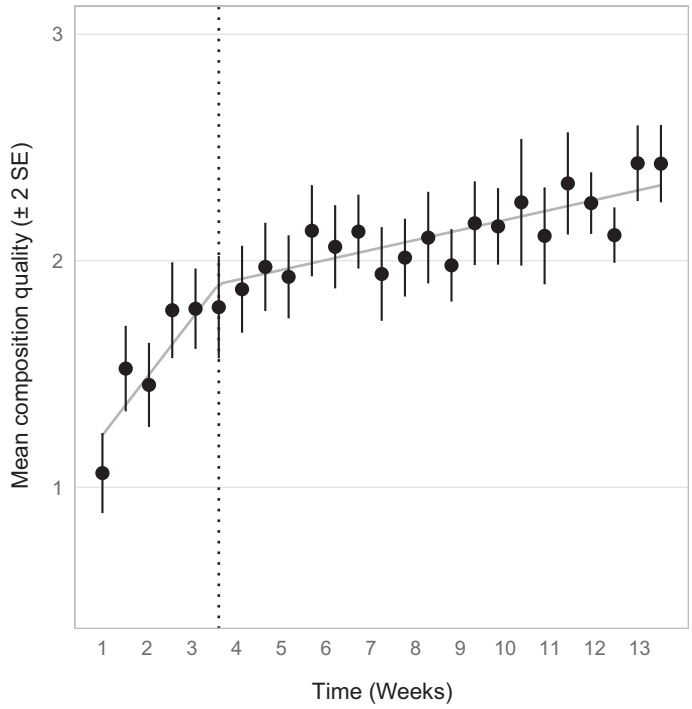

Figure 1. Change in composition quality over time. Points represent observed means at each of 25 composition test occasions. Curves represent growth curve estimates for the first 3.5 weeks and the following 9.5 weeks of the study.

better fit than any less complete alternatives ($\chi^2 > 100, p < .001$ relative to all competing models). Statistical significance for model parameters was evaluated by *z*-test.

Learning rates

We evaluated a series of models, each assuming a different demarcation between initial and later learning rates. We represented fixed effects for test occasion (time from start of school) with separate dummy variables representing linear growth across the initial period (*Period 1*) and across the remainder of the test occasions (*Period 2*). Effects of the two timepieces (periods) were allowed to correlate. We evaluated models with splits ranging from 2nd composition test (i.e., Period 1 slope for just test occasions 1 and 2) through to the 15th test (Period 1 slope for test occasions 1–15). Fits for all of these models can be found in Appendix D. We found best fit where Period 1 ranged over the first six tests and Period 2 ranged over the remaining 19. AIC for this model was at least 10 less than for all other models (strong evidence of better fit; Burnham & Anderson, 2004) with the exception of the model which split at Test 7 (ΔAIC = 3.3). Therefore in subsequent analyses evaluated models with separate growth curves for test occasions 1—6 (first 3.5 weeks of testing) and test occasions 7–25 (final 9.5 weeks).

We then evaluated three incremental models, starting with a baseline (intercept only) model (Model 0), then adding the Period 1 slope (Model 1), and then adding the Period 2 slope (Model 2). Model 1 showed improved fit relative to Model 0 ($\chi^2(1) = 15, p < .001$)

indicating statistically significant increase in learning across the first 3 weeks. Model 2 showed improved fit relative to Model 1 ($\chi^2(1) = 10$, $p = .001$), indicating statistically significant but more gradual increase in learning across the final 9.5 weeks of the study. Growth curve estimates from this final model are shown in Figure 1.

Finally, we determined whether this piecewise model (Model 2) provided better fit to our data than a model in which the two fixed timepiece effects were replaced by a single linear growth curve. We found strong evidence that Model 2 provided better fit (ΔAIC = 22).

Factors affecting learning rates

Table 1 gives bivariate correlations among the various start-of-year measures. As might be expected given that Spanish orthography has very regular grapheme-phoneme correspondence, pseudoword, and real-word spelling scores were strongly correlated. We therefore combined these into a single spelling ability measure. Similarly performance on the accurate and speeded copy tasks was strongly correlated, and we combined these to create a single transcription fluency measure. In both cases, variables were standardized then summed. Alphabet copying, which was included in this study as a potential measure of transcription fluency because of its use in previous research, did not correlate strongly with performance on the sentence-copying tasks and so was dropped from further analysis.

We determined the effects of the predictor variables (age, non-verbal skill, phonological awareness, letter knowledge, single-word reading, handwriting accuracy, transcription fluency, and spelling) on overall written composition performance, on initial learning rate (Period 1 slope), and on subsequent learning rate (Period 2 slope) as follows: We evaluated a sequence of seven incremental models at each stage adding, for one predictor variable, main effect, interaction with the Period 1 timepiece dummy variable, and interaction with the Period 2 timepiece dummy variable. We started with a model with main effects for Period 1 and Period 2 (i.e., Model 2 detailed in the previous section) and added predictor variables, starting with control variables that we hypothesized might be related to composition ability but were not directly implicated in text production (age, non-verbal skill, phonological awareness, letter knowledge, single-word reading; Models 3–7), and then adding handwriting accuracy, transcription fluency and spelling ability (Models 8–10) which we hypothesized as direct, causal predictors of composition performance and of preparedness to learn to compose. Both predictor and dependent variables were standardized prior to analysis. Each subsequent model provided significantly better fit relative to the previous model (Model 3, $\chi^2(3) = 11$, $p = .011$; all other models, $\chi^2(3) > 22$, $p < .001$). The final, best fit model (Model 10) gave a marginal R^2 (Nakagawa & Schielzeth, 2013) of .63. Variance breakdown is shown in Table 2.

Parameter estimates for the final model can be found in Table 3. These show strong evidence for positive main effects on written composition performance (i.e., effects on aggregate performance across all 13 weeks) for non-verbal ability, word reading ability, transcription fluency, and spelling ability, and weaker evidence for a positive effect of handwriting accuracy.

Rate of initial learning (Period 1 slope) was negatively correlated with single-word reading – children with poorer reading showed more rapid initial improvement – but positively correlated with spelling ability. However, later learning (Period 2 slope) was negatively associated with initial spelling ability: Weaker spellers showed faster learning, relative to their peers, in the last 9.5 weeks of the study. We also found much weaker

Table 1. Means and bivariate correlations for predictor variables

	M (*SD*)	Age	Non-verbal	Phon. aware.	Letter know.	Reading	Handwriting	Copy alphabet	Copy accurate	Copy speed	Spell real
Age (months)	74.5 (3.4)										
Non-verbal skill	14.6 (2.9)	.08									
Phonological Awareness	15.4 (10)	.16	.28								
Letter knowledge	24.1 (4.1)	.13	.34	.25							
Word reading	0.00 (5.2)	.07	.21	.49	.15						
Handwriting accuracy	3.03 (.94)	.20	.41	.32	.37	.33					
Copy alphabet	8.94 (3.9)	.14	.06	.33	.19	.17	.22				
Copy accurate	2.83 (2.1)	.16	.07	.22	.32	.27	.24	.18			
Copy speed	3.4 (2.2)	.31	.19	.35	.41	.36	.27	.17	.69		
Spelling real words	12.1 (6.2)	.17	.40	.42	.58	.42	.53	.14	.56	.60	
Spelling pseudowords	12.9 (6.3)	.15	.42	.46	.55	.46	.52	.19	.52	.59	.87

Note. For correlation coefficients (Pearson's *r*), $p < .01$ for $r > |.2|$.

Table 2. Factors affecting learning rate: Variance components for final model (Model 10)

All fixed effects (Table 3)	.4227
Random effects	
Child (nested within classroom)	
Intercept	.0916
Period 1 slope	.0020
Period 2 slope	.0003
Intraclass correlation	.111
Classroom	
Intercept	.0610
Period 1 slope	.0015
Period 2 slope	.0000
Intraclass correlation	.074
Residual	.2663

Table 3. Factors affecting learning rate: Standardized parameter estimates from final model

	Main effect (intercept)	Effect on initial learning rate (Period 1 slope)	Effect on later learning rate (Period 2 slope)
Age	−.004 (.048), .930	.004 (.009), .670	−.001 (.002), .720
Non-verbal skill	.139 (.053), .009	−.011 (.010), .260	−.001 (.002), .790
Phonological awareness	−.017 (.060), .780	.001 (.011), .920	−.000 (.003), .950
Letter knowledge	−.059 (.053), .270	.007 (.010), .460	.006 (.003), .038
Word reading	.180 (.054), .001	−.025 (.010), .014	−.000 (.003), .870
Handwriting accuracy	.116 (.057), .042	.009 (.011), .420	−.003 (.003), .310
Transcription fluency	.357 (.063), .000	−.013 (.012), .270	−.006 (.003), .038
Spelling	.192 (.079), .015	.036 (.015), .017	−.013 (.004), .000

effects on Period 2 slope for letter knowledge (faster learning for higher-scoring students) and for transcription fluency (faster learning for students with lower initial fluency, following the same patter and for spelling). All other effects were either very weak or non-significant.

Spelling effects are illustrated in Figure 2. Students who started the study with typical or good spelling skill tended to show little or no improvement after the first 3.5 weeks of testing. However, students with low initial spelling ability started with weaker performance but improved steadily throughout the study.

Discussion

Our findings can be summarized as follows: (1) Children's written composition performance averaged across the first 13 weeks of their first year at primary (elementary) school was positively correlated with performance on a non-verbal reasoning task, single-word reading, spelling ability, handwriting neatness, and, particularly, transcription (sentence-copying fluency). We did not find evidence for direct, independent effects of age, phonological awareness, or letter knowledge; (2) overall students' composition performance improved rapidly over the first 3.5 weeks of the study and then showed much slower, though still statistically significant growth over the following 9.5 weeks. (3) However, students entering school with weak spelling ability showed a different pattern,

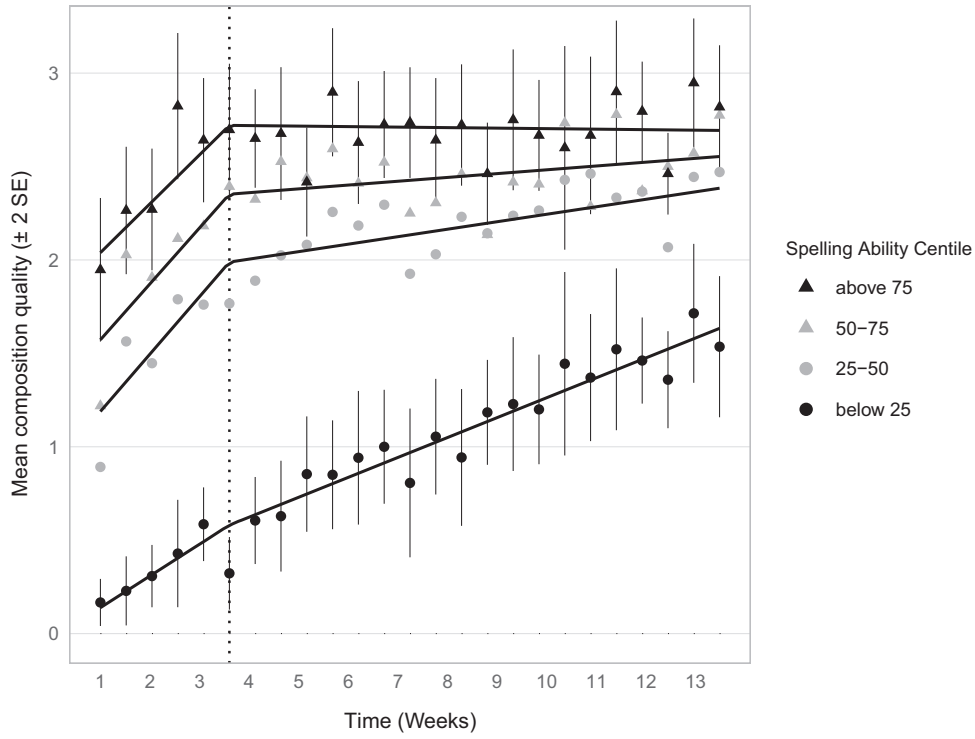

Figure 2. Change in composition quality over time by spelling ability. Points represent observed means at each of 25 composition test occasions, plotted separately for students grouped by spelling ability quartile. For clarity, error bars are omitted for students in the middle two quartiles. Curves represent growth curve estimates for the first 3.5 weeks and the following 9.5 weeks of the study, again plotted separately for each spelling-quartile group.

with initially slower learning then steady improvement through the remainder of the testing period. We will discuss each of these sets of findings in turn.

As we have noted, factors affecting written composition performance in 1st grade, and particularly the first semester of first grade, are interesting because this is, in most educational systems and in the present study, when students first start learning in earnest to produce text. Looking first at main effects on performance, our findings differ to some extent from those of the only previous study to look at the effects of transcription fluency, reading, and spelling, as separate factors, on first-grade writing. Kim et al. (2013) found effects for reading and fluency, but not spelling. We found effects for all three of these factors. One straightforward explanation for this difference may be that our study afforded substantially greater statistical power by dint of our use of multiple performance probes. However, our study differed from theirs in other ways including choice of measures, language, and instructional context and this makes direct comparison difficult. We return to the issue of instructional context below.

Our finding that fluency, spelling, and reading measures – respectively, sentence-copying speed, accurate spelling of both regular and tricky words, and single-word reading for comprehension – correlated with composition quality is, however, consistent with the broader writing development literature summarized in our introduction. Competition for a shared, general working memory is often invoked as explanation for

why transcription ability contributes to text quality (e.g., Berninger, Abbott, Abbott, Graham, & Richards, 2002), although arguably these accounts lack explanatory power (Torrance & Galbraith, 2006). A better understanding, again arguably, is that when central processing is required for orthographic and motor planning, then this disrupts the fluent parallel and cascaded processing that is typical of developed writing (Bonin, Roux, Barry, & Canell, 2012; Olive, 2014): Written production, like speech, is a 'just in time' system (Christiansen & Chater, 2016) in which delays at output result in disruption further upstream. A more parsimonious explanation for these effects, however, particularly given that students were sampled right at the start of their school careers, may simply be that students with low sentence-copying speed and spelling ability lacked the handwriting and spelling knowledge necessary to express their ideas. Children with very weak spelling and handwriting are simply not going to be able to form sufficient words on the page to construct a narrative.

We also found a relatively strong unique contribution of reading. Although reading plays a role in the revision processes of advanced writers producing extended texts, it is probably of less value for simpler tasks (Olive & Piolat, 2002; Torrance, Rønneberg, Johansson, & Uppstad, 2016). Revision is very unlikely to be essential to the quality of very early writers' compositions. We therefore suspect that effects of reading ability in our present study have more to do with ability to fluently map between semantics and orthography. The reading task of course involved rapid mapping of text onto meaning; the reverse of the processing required when writing. However underlying representations are likely to be similar in both cases, and so, fluency in one direction will correlate with fluency in the other.

The particular contribution of our research was that we went beyond predicting composition quality at a single time-point to examine how performance developed over time. We found rapid learning across the first 3.5 weeks of the study, followed by much slower learning. Composition-focussed teaching in the first 3 weeks of the study aimed largely to motivate students, with content focussed on the general importance of learning to write and on finding good ideas to write about. Initial improvement may have resulted directly from instruction improving student motivation. It may also be that, independent of instruction, there were large initial gains from repeating the assessment task as students got used to sustaining attention on a single writing task. Relatively, slow growth after this initial rapid improvement has two possible general explanations. Anecdotally and perhaps predictably, students' motivation waned across multiple tasks. The task – writing a narrative about events in their own lives – was chosen deliberately so that students would always have something to write about. However, repeatedly asking children to find something narrative-worthy from their own experience proved demotivating over time. Therefore, positive effects of instruction may have been offset by negative effects of testing. Alternatively, it may be that for the majority of students instruction was relatively ineffective.

The one clear exception to this pattern was students who entered school with relatively poor spelling skills. These students performed less well than their peers initially but then showed steady improvement throughout the 13 weeks of assessment. (Students with lower transcription fluency showed a similar though much weaker pattern). This effect is not surprising. For reasons suggested above, spelling is implicated in composition quality, even when quality ratings do not include assessment of spelling accuracy. This effect will be thresholded, however. Once a student's spelling ability reaches a level such

that they can, without too much difficulty, generate expressions of simple concepts as accurate or at least readable written words, then their spelling ability will not constrain their ability to produce text that is recognizable as a narrative. In a shallow orthography, this threshold is likely to be quite low. A straightforward explanation for the spelling effect shown in Figure 2 is, therefore, that most students entered school with sufficient spelling ability to create a simple narrative, but a minority of students did not have these skills on entry and acquired them gradually across the 13 weeks of our study. This account is consistent with findings from Juel et al. (1986). They found that in first grade, spelling strongly predicted narrative quality, with a relatively weak effect for ideation. In second grade, after spelling skills had developed, this pattern reversed.

The specific findings of this study must be interpreted in the context of its specific instructional and language context. Writing, unlike speech, is acquired through direct instruction, and studies that explore writing performance are necessarily examining students' response to that instruction. This is true whether students' performance is captured at a single time-point or as it changes over time. For example, the failure of Kim et al. (2013) to find evidence for effects of spelling ability, in contrast to present findings, was in the context of a study conducted later in first grade following differential instruction with targeted support for weaker learners. This is a no less valid context in which to explore predictors of composition performance than the present study. It simply demonstrates that writing performance and how this changes over time is a response to intervention rather than a more general cognitive-developmental trend.

Reviewers of the first version of this paper raised two further important issues. First, the measures used in this study were, in all cases, developed or adapted specifically for this study. In all cases, our measures have good internal reliability and, we believe, good face validity. However, use of established measures, had appropriate Spanish language tools been available, may have made direct comparison with previous studies more straightforward. We have, however, provided detailed description of the tools that we used to permit easy comparison with the (wide range of) measures in the previous literature. Second, in our present study the quality of written composition was scored in terms of a single, holistic rating. Some previous studies (e.g., Kim et al., 2013; Wagner et al., 2011) differentiated between accuracy of spelling and handwriting, and higher level featured of the text (idea development, organization, and so forth). A similar distinction in the present study would have been helpful in unpicking the mechanisms underlying, for example, the effect of start of year spelling test performance and their subsequent written composition learning.

Conclusion

The present study is one of very few to explore predictors of written composition quality in students' first year at school, and the only study that we are aware of that has systematically explored effects of these predictors on student learning over time. In the context of a specific programme of instruction that focussed both on transcription (handwriting, spelling) and ideation (text planning), we found that first-grade students who start school with weaker transcription skills relative to peers produce weaker written narratives but showed greater improvement in response to instruction. We do not make strong claims about the generalizability of these findings. However, they do point towards the potential value of differential instruction when teaching students to compose text, from the start of school, and of subsequent close monitoring of response to intervention.

Acknowledgements

This research was funded by the Spanish Ministry of Economy and Competitiveness [Ministerio de Economía y Competitividad de España], grant EDU2015-67484-P MINECO/FEDER. The funders had no direct input into the design and implementation of the research. María Arrimada benefited from a research grant (FPU14/04467) awarded by the Spanish Ministry of Education, Culture and Sports (Ministerio de Educación, Cultura y Deporte de España). Mark Torrance is part-funded by the Norwegian Centre for Reading Education and Research, University of Stavanger, Norway. A small subsample of data reported in this paper was analysed and reported for different purposes in Arrimada et al. (2018a).

Conflicts of interest

All authors declare no conflict of interest.

Author contribution

Mark Torrance (Conceptualization; Formal analysis; Methodology; Supervision; Visualization; Writing – original draft; Writing – review & editing) María Arrimada (Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Writing – original draft; Writing – review & editing) Sarah Gardner (Formal analysis; Writing – original draft; Writing – review & editing).

Data availability statement

Data from this study are available at https://doi.org/10.17605/OSF.IO/3U5ZD

References

Abbott, R.D., & Berninger, V.W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology*, *85*, 478–508. https://doi.org/10.1037/0022-0663.85.3.478

Arrimada, M., Torrance, M., & Fidalgo, R. (2018a). Supporting first-grade writers who fail to learn: Multiple single-case evaluation of a Response to Intervention approach. *Reading and Writing*, *31*(4), 865–891. https://doi.org/10.1007/s11145-018-9817-x

Arrimada, M., Torrance, M., & Fidalgo, R. (2018b). Effects of teaching planning strategies to first-grade writers. *British Journal of Educational Psychology*, 1–19. https://doi.org/10.1111/bjep.12251

Babayiğit, S., & Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: A longitudinal study. *Reading and Writing*, *23*, 539–568. https://doi.org/10.1007/s11145-009-9173-y

Babayiğit, S., & Stainthorp, R. (2011). Modeling the relationships between cognitive – Linguistic skills and literacy skills. *New Insights from a Transparent Orthography.*, *103*(1), 169–189. https://doi.org/10.1037/a0021671

Barnett, A., Henderson, S.E., Scheib, B., & Schulz, J. (2009). Development and standardization of a new handwriting speed test: The Detailed Assessment of Speed of Handwriting. *British Journal of Educational Psychology*, *2*(6), 137–157. https://doi.org/10.1348/000709909X421937

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berninger, V.W., Abbott, R.D., Abbott, S.P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities*, *35*(1), 39–56. https://doi.org/10.1177/002221940203500104

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). Guildford Press.

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. https://doi.org/10.1007/BF01027151

Binamé, F., & Poncelet, M. (2016). Order short-term memory capacity predicts nonword reading and spelling in first and second grade. *Reading and Writing*, *29*(1), 1–20. https://doi.org/10.1007/s11145-015-9577-9

Blatchford, P. (1991). Children's writing at 7 years: Associations with handwriting on school entry and pre-school factors. *British Journal of Educational Psychology*, *61*(1), 73–84. https://doi.org/10.1111/j.2044-8279.1991.tb00962.x

Bonin, P., Roux, S., Barry, C., & Canell, L. (2012). Evidence for a limited-cascading account of written word naming. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(6), 1741–1758. https://doi.org/10.1037/a0028471

Burnham, K.P., & Anderson, D.R. (2004). Multimodel inference. *Sociological Methods and Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Chou, C.P., Yang, D., Pentz, M.A., & Hser, Y.I. (2004). Piecewise growth curve modeling approach for longitudinal prevention study. *Computational Statistics and Data Analysis*, *46*(2), 213–225. https://doi.org/10.1016/S0167-9473(03)00149-X

Christiansen, M.H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, 1–72. https://doi.org/10.1017/S0140525X1500031X

Coker, D. (2006). Impact of First-grade factors on the growth and outcomes of urban schoolchildren's primary-grade writing. *Journal of Educational Psychology*, *98*, 471–488. https://doi.org/10.1037/0022-0663.98.3.471

Cornhill, H., & Case-Smith, J. (1996). Factors that relate to good and poor handwriting. *American Journal of Occupational Therapy*, *50*, 732–739. https://doi.org/10.5014/ajot.50.9.732

Daly, C.J., Kelley, G.T., & Krauss, A. (1994). Relationship between visual-motor integration and handwriting skills of children in kindergarden: A modified replication study. *American Journal Occupational Therapy*, *57*, 459–462. https://doi.org/10.5014/ajot.57.4.459

Defior, S., Jiménez-Fernández, G., & Serrano, F. (2009). Complexity and lexicality effects on the acquisition of Spanish spelling. *Learning and Instruction*, *19*(1), 55–65. https://doi.org/10.1016/j.learninstruc.2008.01.005

Dunsmuir, S., & Blatchford, P. (2004). Predictors of writing competence in 4- to 7-year-old children. *The British Journal of Educational Psychology*, *74*, 461–483. https://doi.org/10.1348/0007099041552323

Frolek, G., & Luze, G. (2014). Predicting handwriting performance in kindergarteners using reading, fine-motor, and visual-motor measures. *Journal of Occupational Therapy, Schools, and Early Intervention*, *7*(1), 29–44. https://doi.org/10.1080/19411243.2014.898470

Frost, J. (2001). Phonemic awareness, spontaneous writing, and reading and spelling development from a preventive perspective. *Reading and Writing*, *14*, 487–513. https://doi.org/10.1023/A:1011143002068.

Jiménez, J.E., & Hernández-Cabrera, J.A. (2019). Transcription skills and written composition in Spanish beginning writers: Pen and keyboard modes. *Reading and Writing*, *32*, 1847–1879. https://doi.org/10.1007/s11145-018-9928-4

Jones, D., & Christensen, C.A. (1999). Relationship between automaticity in handwriting and students' ability to generate written text. *Journal of Educational Psychology*, *91*(1), 44–49. https://doi.org/10.1037//0022-0663.91.1.44

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*, 437–447. https://doi.org/10.1037/0022-0663.80.4.437

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. https://doi.org/10.1037/0022-0663.78.4.243

Kent, S.C., & Wanzek, J. (2016). The Relationship between component skills and writing quality and production across developmental levels. *Review of Educational Research*, *86*(2), 570–601. https://doi.org/10.3102/0034654315619491

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. https://doi.org/10.1007/s11145-013-9480-1

Kim, Y.-S., Al Otaiba, S., Folsom, J., & Gruelich, L. (2013). Language, literacy, attentional behaviors and instructional quality predictors of written composition for first graders. *Early Childhood Research Quarterly*, *28*(3), 461–469. https://doi.org/10.1016/j.ecresq.2013.01.001

Kim, Y.-S., Al Otaiba, S., Wanzek, J., & Gatlin, B. (2015). Towards an understanding of dimensions, predictors, and gender gap in written composition. *Journal of Educational Psychology*, *107*(1), 79–95. https://doi.org/10.1037/a0037210

Kim, Y.S., Otaiba, S.A., Folsom, J.S., Greulich, L., & Puranike, C. (2014). Evaluating the dimensionality of first-grade written composition. *Journal of Speech, Language, and Hearing Research*, *57*(1), 199–211. https://doi.org/10.1044/1092-4388(2013/12-0152)

Kim, Y.S.G., & Park, S.H. (2019). Unpacking pathways using the direct and indirect effects model of writing (DIEW) and the contributions of higher order cognitive skills to writing. *Reading and Writing*, *32*, 1319–1343. https://doi.org/10.1007/s11145-018-9913-y

Lehtonen, A., & Bryant, P. (2004). Length awareness predicts spelling skills in Finnish. *Reading and Writing*, *17*(9), 875–890. https://doi.org/10.1007/s11145-004-2802-6

Mäki, H.S., Voeten, M.J.M., Vauras, M.M.S., & Poskiparta, E.H. (2001). Predicting writing skill development with word recognition and preschool readiness skills. *Reading and Writing*, *14*, 643–672. https://doi.org/10.1023/A:1012071514719

Martínez, J., & García, E. (2004). *Diccionario: Frecuencias del castellano escrito en niños de 6 a 12 años [Dictionary: Spanish frequencies written in children aged 6 to 12 years.]*, Publicaciones de la Universidad Pontificia.

McCutchen, D. (1996). A capacity theory of writing: Workin memory in composition. *Educational Psychology Review*, *8*(3), 299–325. https://doi.org/10.1007/bf01464076

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R 2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly*, *32*(2), 154–167. https://doi.org/10.1598/rrq.32.2.2

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

Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*, *6*(2), 173–194. https://doi.org/10.17239/jowr-2014.06.02.4

Olive, T., & Piolat, A. (2002). Suppressing visual feedback in written composition: Effects on processing demands and coordination of the writing processes. *International Journal of Psychology*, *37*, 209–218. https://doi.org/10.1080/00207590244000089

Puranik, C.S., Lonigan, C.J., & Kim, Y.-S. (2011). Contributions of emergent literacy skills to name writing, letter writing, and spelling in preschool children. *Early Childhood Research Quarterly*, *26*, 465–474. https://doi.org/10.1016/j.ecresq.2011.03.002

Raven, J.C. (1981). *Manual for Raven's progressive matrices and vocabulary scale*, : . Psychological Corporation.

Rijlaarsdam, G., & Couzijn, M. (2000). Writing and learning to write: A double challenge. In R. Simons, J. van der Linden & T. Duffy (Eds.), *New learning* (pp. 157–189). Springer. https://doi.org/10.1007/0-306-47614-2_9

Sadoski, M., Willson, V.L., Holcomb, A., & Boulware-Gooden, R. (2004). Verbal and nonverbal predictors of spelling performance. *Journal of Literacy Research*, *36*, 461–478. https://doi.org/10.1207/s15548430jlr3604_2

Taylor, J., & Schatschneider, C. (2010). Genetic influence on literacy constructs in kindergarten and first grade: Evidence from a diverse twin sample. *Behavior Genetics*, *40*, 591–602. https://doi.org/10.1007/s10519-010-9368-7

Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C.A. MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 67–82). Guildford Publications.

Torrance, M., Rønneberg, V., Johansson, C., & Uppstad, P.H. (2016). Adolescent weak decoders writing in a shallow orthography: Process and product. *Scientific Studies of Reading*, *20*, 375–388. https://doi.org/10.1080/10888438.2016.1205071

Tseng, M.H., & Murray, E.A. (1994). Differences in perceptual-motor measures in children with good and poor handwriting. *The Occupational Therapy Journal of Research*, *14*(1), 19–36. https://doi.org/10.1177/153944929401400102

van den Bergh, H., Maeyer, S., van Weijen, D., & Tillema, M. (2012). Generalizability of text quality scores. In E. Van Steendam, M. Tillema, G. Rijlaarsdam, & H. van den Bergh (Eds.), *Measuring writing: Recent insights into theory, methodology and practice* (pp. 23–32). Brill. https://doi.org/10.1163/9789004248489_003

Van Steensel, R. (2006). Relations between socio-cultural factors, the home literacy environment and children's literacy development in the first years of primary education. *Journal of Research in Reading*, *29*, 367–382. https://doi.org/10.1111/j.1467-9817.2006.00301.x

Wagner, R.K., Puranik, C.S., Foorman, B., Foster, E., Wilson, L.G., Tschinkel, E., & Kantor, P.T. (2011). Modeling the development of written language. *Reading and Writing*, *24*(2), 203–220. https://doi.org/10.1007/s11145-010-9266-7

Received 8 February 2020; revised version received 5 November 2020

Appendix A: Spelling task

Real words

Words were chosen to include one of the following features: *a complex grapheme* – a phoneme represented by two letters written together; *a contextual effect* – a specific consonant represents different phonemes depending on the vowel accompanying it; *a position effect* – a consonant represents different phonemes depending on its position within the word; *an inconsistency* – a phoneme that can be represented by two or more graphemes without a specific rule to determine which is correct; *letter H*, which is a silent letter in Spanish, and there are no specific rules to place it correctly within a word; and *stress mark*.

The words used were as follows: paquete, guitarra, acera, cisne, ramo, carroza, general, jirafa, hada, hechizo, camión, colocó (*package, guitar, pavement, swan, bouquet, carriage, general, jiraffe, fairy, spell, truck, put*).

Pseudowords

taquimo, guesirre, ociro, celto, ruca, corrizo, giteros, gerraso, hepo, haqueza, cuseón, caticú.

Appendix B: Handwriting assessment criteria

Score	Criteria
0	The majority of the marks in the paper do not resemble letters. They can be pictures or random strokes. There might be some strokes that resemble letters but most of them don't.
1	The majority of the marks on the first line can be identified as specific letters (independent of context). However nearly all of these are very badly formed in one or more ways: they are inaccurate (i.e., you can hardly tell which letter the strokes correspond to) and irregular (i.e., shaky strokes, different sizes, same size for capitals and non-capitals, oscillations, slant letters, and letters overlapped).
2	The majority of the marks on the first can be identified as specific letters (independent of context). The majority of these are accurately formed but are irregular. Irregularity can include (see above). Letter size is consistent throughout the text but they are so small it is difficult to know whether they are accurate or not.
3	The majority of the marks on the first line can be identified as specific letters (independent of context). At least half of the letters are accurate and regular.
4	All the marks on the first line can be identified as specific letters (there can be 1 mark not identified as a letter). The great majority of letters are accurate and regular. There might be some irregularities.

Appendix C: Composition quality assessment criteria

Score	Criteria
0	There is no text or it is illegible. The text does not respond to the topic. The text is a list of words without clauses.
1	Certain progression of ideas: the child mentions 1 or 2 clauses without clarifying or descriptive details. Frequent digressions. Handwriting is difficult to understand, and there are frequent spelling mistakes. Grammar complexity: simple sentences. No connectors or very repetitive ones. Basic and simple vocabulary.
2	Certain progression of ideas with one of the following: 1 or 2 ideas mentioned with descriptive or clarifying details. More than 2 ideas with very few or no descriptive or clarifying details. Repetitive and irrelevant details. Legible handwriting with common spelling mistakes. Grammar complexity: mostly simple sentences but there are some compound ones formed by juxtaposition (connector 'and'). Basic and repetitive connectors. Vocabulary typical for students' age.
3	Logical progression/sequence of ideas, linked to a common topic and with descriptive and clarifying details. Some irrelevant or repetitive details. Correct handwriting with some spelling mistakes. Grammar complexity: combination of simple and compound sentences (mostly juxtaposition but some formed by subordination). Basic and repetitive connectors, although they might include a complex one. Appropriate vocabulary.
4	Logical progression/sequence of ideas, linked to a common topic and with a variety of descriptive and clarifying details. No irrelevant details although there might be some repetitions. Correct handwriting with some spelling mistakes. Grammar complexity: mostly compound sentences (juxtaposition, coordination and subordination) Repetitive connectors although some complex ones. Appropriate vocabulary with a few unusual expressions/words.
5	Logical progression/sequence of ideas, linked to a common topic and with a variety of descriptive and clarifying details. Variety of relevant and non-repetitive details. Correct handwriting with very few spelling mistakes. Grammar complexity: mostly compound sentences. Varied connectors. Advanced vocabulary. Certain textual structure: introduction, development, and conclusion.

Appendix D: Model fits for evaluation of different timepiece demarcation locations

The following table gives goodness of fit statistics for models predicting writing performance on the basis on test occasion divided into two linear growth curves representing initial tests and later tests. Models varied in the test occasion that represented the boundary between the two timepieces. *Demarcation* represents that last test occasion of the first timepiece. The best fit model put this at the 6th test occasion. *AIC difference* is the difference in AIC relative to this model.

Demarcation	AIC	AIC difference
2	6,279.5	91.7
3	6,247.9	60.1
4	6,212.3	24.6
5	6,198.6	10.9
6	6,187.8	
7	6,191.7	3.94
8	6,204.2	16.5
9	6,217.9	30.1
10	6,230.9	43.2
11	6,256.6	68.9
12	6,277.1	89.3
13	6,294.6	106
14	6,308.1	120
15	6,315.9	128