An Investigation of the Interaction between Writing Modality and Literacy Skills in Children Learning Written Composition

Eivor Finset Spilling[1], Vibeke Rønneberg[2], Wenke Mork Rogne[1], Jens Roeser[3] & Mark Torrance[3]

[1]Volda University College | Norway

[2]University of Stavanger | Norway

[3]Nottingham Trent University | UK

Abstract: Start-of-school writing instruction traditionally entails students writing by hand, but alternatively can involve keyboarding, with handwriting instruction delayed. We tested the hypothesis that which approach most benefits a specific child is dependent, in part, on their literacy skills at school entry. In an educational context in which minimal literacy instruction occurs prior to first grade, we compared five first-grade classes in which students (N = 88) learned by hand with five classes in which students (N = 89) learned to write by typing on digital tablets. Students completed tasks measuring phoneme isolation, phoneme blending, grapheme-to-phoneme mapping, word reading, spelling and vocabulary at school entry, and composed narrative texts at five timepoints spread across the school year. The present study reports an additional analysis of data first reported in Spilling et al. (2023). Bayesian statistical modelling established effects of start-of-school literacy on overall composition performance, but not on students' rate of learning. We found moderate to strong evidence against the hypothesis that start-of-school literacy affected rate of learning differently depending on the modality in which children were taught: Neither writing by pencil and paper nor writing by keyboard provides specific benefits for first-grade students with lower – or higher – literacy skills learning to compose text.

Keywords: handwriting, digital tablet, literacy effects, modality effects, written composition

journal of
WRITING RESEARCH

Spilling, E. F., Rønneberg, V., Rogne, W. M., Roeser, J. & Torrance, M. (Accepted for publication in 2026). An investigation of the interaction between writing modality and literacy skills in children learning written composition. *Journal of Writing Research, volume(issue),* ##-##. DOI: xx

Contact: Eivor Finset Spilling, Department of Language and Literature, Volda University College, Mailbox 500, 6101 Volda | Norway – eivor.finset.spilling@hivolda.no

1. Introduction

School starters have different points of departure for learning to write. Exposure to formal and informal literacy instruction prior to starting school varies substantially from child to child resulting in substantial variation in learning (Whitehurst & Lonigan, 1998). In a group of first graders some children will be able to identify many letters, map sounds and letters, and even spell and decode words. In contrast, others will have limited phonological awareness and knowledge of letters. Early literacy skills are related to writing performance and to learning to write by hand (Kim et al., 2014; Torrance et al., 2021).

The writing modality used for initial writing instruction also varies. In most educational systems children first learn to write using pencil and paper. However, some schools teach writing with keyboard from the start of first grade (Genlott & Grölund, 2013). The backdrop is the increasing digitisation of educational systems. Writing by keyboard has been included in national curricula even in early grades in several educational contexts, for example in Nordic and English-speaking countries (Gamlem et al., 2020; Malpique et al., 2023a; Poole & Preciado, 2016). Furthermore, there has been a trend worldwide for schools to supply students with personal digital devices for school work (Islam & Grönlund, 2016). In Norway, there was a wave of providing students with personal digital devices starting in around 2018. Within a few years a majority of schools provided students with digital tools, with 80% of students in first to fourth grade having access to their own digital device, typically digital tablet, in school (The Norwegian Directorate for Education and Training, 2022). Digital tablets are still the most commonly used digital writing tool in these early grades in Norway.

Writing modality can affect writing performance and learning to write. However, findings are inconsistent. A recent meta-analysis of students in kindergarten through Grade 6 (Malpique et al., 2023c) found positive effects of handwriting on composition quality and mixed effects on letter production and word production, with younger students producing more letters/words by hand and older students writing more by keyboard. Goldberg et al.'s (2003) meta-analysis of students in kindergarten through Grade 12, on the other hand, showed a positive effect of keyboarding on writing quality, with smaller effects for primary students. The inconsistent findings of modality effects can be related to what writing skills are targeted, and also how used participants are to the writing modalities.

1.1 Literacy skills

Learning written composition involves learning how to generate and structure ideas and then to translate these ideas into handwriting or print (Berninger & Winn, 2006; Juel et al., 1986). The underlying abilities to translate ideas into text are related to phonological awareness, letter–sound knowledge, linguistic and lexical knowledge, and knowledge of print (Juel et al., 1986; Whitehurst & Lonigan, 1998). The development of these skills starts long before the formal literacy teaching in school, and many preschool children will have begun to make attempts at writing, such as forming letters, writing their own name, and spelling simple words either with invented spelling or conventional spelling (Albuquerque & Alves Martins, 2022;

Puranik et al., 2011). The extent to which these literacy skills are developed in school starters, will vary depending on home support for literacy skills and pre-school instruction and language stimulation. Therefore, first-grade children will vary considerably in the literacy skill that they bring with them to start of school and to the task of learning to compose text.

Several early literacy skills have been identified as predictors of first-grade children's writing performance when writing by pencil and paper. Both writing productivity and quality, and – though scarcely explored – children's rate of learning to write are related to literacy skills. Correlational studies give evidence that the ability to read, both single-word decoding and reading comprehension, is related to the number of words, sentences, and ideas produced (Berninger et al., 2002; Kim et al., 2014) and to quality (Berninger et al., 2002; Kim et al., 2013; Kim et al., 2014). There is also evidence that performance on single-word spelling-to-dictation tasks has an effect on text quality (Jiménez & Hernández-Cabrera, 2018; Juel et al., 1986). Similarly, vocabulary and grammatical knowledge are related to quality (Kim et al., 2013; Kim et al., 2014). The studies cited above sampled students writing by hand. Very few studies have focussed on children writing digitally, but Malpique et al. (2023a) investigated student-level effects on written composition fluency and quality in a sample of 544 second graders composing by keyboard. They found that spelling-to-dictation predicted quality and that word-reading ability predicted composition fluency and quality.

Longitudinal studies also show that there are relations between children's spelling, reading, and oral language skills and writing performance by hand at subsequent time-points. Kent et al. (2014) found that a composite measure of spelling and reading and a composite measure of vocabulary and grammatical knowledge in kindergarten predicted compositional quality in first grade. Coker (2006) found that reading (identification of letters and words) and receptive vocabulary measured at school start predicted writing quality in descriptive texts at the end of first grade. Both Dunsmuir and Blatchford (2004) and Babayiğit and Stainthorp (2011) demonstrated that students' vocabulary measured in first grade predicted the quality of texts written in the second grade. Further, first-grade spelling ability has been found to predict compositional coherence (Mäki et al., 2001) and compositional spelling and word length (Oddsdóttir et al., 2021) in texts written by second graders.

One study, by Torrance et al. (2021), investigated how children's literacy skills measured at school start affect their subsequent rate of learning to compose narratives in Spanish students ($N = 179$) writing by pencil and paper. Results showed that spelling ability, word reading ability, and letter knowledge influenced how fast children learned to compose text. Most students showed rapid initial learning, while students with poor spelling ability showed slow initial improvement and then faster growth compared to peers in the subsequent period. Word-reading ability was negatively correlated to growth in the first period, indicating that poor readers initially showed more rapid growth, while letter knowledge was positively correlated to learning rate both in the first and the last period. This study illustrates that in a first-grade classroom how children respond to instruction is, in part, dependent on their initial literacy skills. The authors point out that since writing is acquired through instruction, effects of literacy skills on learning to write are likely to be specific to the instructional context.

1.2 Modality

Another strand of research compares children's writing performance when composing by keyboard and with pencil and paper. Several studies (Berninger et al., 2009; Malpique et al., 2023b; Read, 2007) have found positive effects of handwriting on quality and length compared to keyboarding in second graders. However, in these studies children were more used to handwriting.

When children receive equal instruction in both modalities, there is evidence that these effects disappear. Spilling et al. (2022) found no modality effects in a sample of first graders receiving instruction in both handwriting and keyboarding from the start of school on quality and length.

There are also studies that have looked at modality effects on learning writing. For the very youngest children, i.e. aged 4 to 6 years, some studies show that learning letters by pencil and paper lead to better letter recognition compared to learning by keyboarding (Longcamp et al., 2005; Mayer et al., 2020), while other studies find no modality effects (Duiser et al., 2022; Kiefer et al., 2015). Some studies have sampled the youngest school children and studied modality effects on text writing performance over time. In an early study Jones and Pellegrini (1996) found that over a ten-week training period, first graders wrote narratives with greater lexical density and grammatical and lexical cohesion when typing, with available text-to-speech support, compared to when they wrote by hand. Genlott and Grönlund (2016) found no differences in scores on national literacy tests in third grade between Swedish students who from the start of school either were given traditional writing instruction or used personal digital devices. Rogne and Rønneberg et al. (2024) investigated the effect of the degree of digitalisation on first graders' writing and found that after the first year of schooling students who were more experienced with digital writing, wrote longer and better-spelled text when writing on a digital device, while students who had done more writing with pencil and paper performed better on handwriting measures. None of these studies, however, investigated performance tracked over time with students writing several texts in the modality they use to learn writing.

1.3 Literacy skills may moderate writing modality effects

A question that remains unanswered is whether the effect of modality on writing might be moderated by students' literacy skills. To the best of our knowledge, only one study, Spilling et al. (2022, also referred to in 1.2), has investigated whether children's literacy skills moderate the effect of writing modality on writing performance. This study found moderate to strong evidence against this being the case. As far as we are aware, however, no study has explored whether there is an interaction effect between modality and children's initial (start-of-school) literacy skills on children's learning of text composition.

Regardless of writing modality, learning to compose text requires repeatedly writing text, and in a particular writing situation, novice writers face a double challenge – producing text and, at the same time, developing their writing skills (Rijlaarsdam & Couzijn, 2000). During transcription, writing by pencil and paper and by keyboard differ in cognitive demands. When handwriting, the writer has to retrieve graphemes from memory and select the correct allographs (van Galen, 1991). When typing, students have access to visual representations of the letters. Thus, typing relies more on recognition of the target letter, and ability to differentiate it from multiple competitors, and less on direct retrieval of the letter form and/or associated motor programs. This may be less demanding at particular stages of writing development.

Writing digitally can also provide support for spelling, via spell checking or text-to-speech synthesis. Even if a child has good ideas, if they cannot spell the words, the ideas will never make it to the page. With the support provided by digital tools children with low spelling skills might be able to produce more text. Additionally, with more text produced the teacher might give them more feedback from which they can learn even more. Studies indicate that teacher feedback on primary-grade children's texts focuses on low-level skills, like handwriting and spelling (Fiskerstrand & Gamlem, 2024; Lucero et al., 2018; Lunsford & Lunsford, 2008; Rønneberg & Nilsen, 2022). Thus, if handwritten texts typically have more errors, one could expect teachers to comment on handwriting and spelling more than when giving feedback to digital texts, while for digital texts there is more scope for commenting on content, thus supporting composition. Further, if students struggle more with spelling when writing by hand, they might forget ideas they were going to write because they focus on correct spelling. There is, however, mixed evidence that this is the case (Rønneberg et al., 2022). If it is true that disruption during writing can influence text quality and that children struggle more with spelling when writing by hand, children with low literacy skills would benefit from learning to write by keyboard on a digital device.

Moreover, the text-to-speech synthesis provided by digital devices, which typically is used as spelling support for young writers, might function better for correcting spelling errors in shorter and more regular words as the child has to interpret the feedback and identify the errors. This means that children with a rich vocabulary might not benefit from digital support the same way as children with low literacy skills in correcting spelling errors.

In contrast, children with well-developed literacy skills might benefit from writing by hand, as letter knowledge is associated with more fluent handwriting (Fitjar et al., 2021). It remains unclear whether the same is true for keyboarding. This means that children with high literacy skills might be able to exploit these skills better when learning to compose text by hand.

1.4 The present study

In the present study we test the hypothesis that writing by hand or digitally when learning to compose text, will benefit children differently depending on their literacy skill at school start. We report additional analysis of data first reported in Spilling et al. (2023), who determined the effects of modality on learning written composition in a sample of 181 Norwegian first graders, where half of the group learned to write by pencil and paper and the other half learned to write by keyboard on a digital tablet. Written composition was operationalised as a series of text features capturing both transcription and ideation, and learning rates for these features, i.e. rates of composition improvement over the first school year, were estimated.

Results showed that the learning rates for text length, syntactic complexity, and narrative sophistication were similar in both modalities. For transcription accuracy, on the other hand, that is, spacing, spelling, and terminator accuracy, students writing by keyboard started at a higher performance level, but experienced little or no improvement compared to students writing by hand. That is, the main finding from Spilling et al. (2023) is a lack of modality effect with the exception of spelling, spacing and terminator accuracy. However, individual differences in literacy skills were not taken into account. In the present study we add data on students' literacy skills to determine if the learning rates in the two modalities found in Spilling et al. (2023) can be observed regardless of the literacy skills that children bring to school. To test our hypothesis, we first determine the effects of literacy skills on composition performance and changes in composition performance through first grade. Then we investigate whether these effects interact with modality on children's learning of text composition the first school year. If modality differentially affects children's learning depending on their literacy skills at school start, we will find different learning rates for students with similar literacy skills learning writing in different modalities.

The statistical modelling was carried out using Bayesian methods. For the present study this approach offers two advantages (cf. Kruschke, 2015; Lambert, 2018; Nicenboim & Vasishth, 2016). Bayesian methods allow one not only to demonstrate evidence in favour of the alternative hypothesis but also in favour of the null hypothesis (Dienes, 2014). Second, frequentist mixed-effects models often fail to converge, particularly for complex models of the form tested in the present study. Use of Bayesian methods permitted full flexibility in how we modelled our data, without concerns over model convergence (Bates et al., 2015).

2. Methods

2.1 Design and participants

The present study is a natural experiment in which ten first-grade classes from ten state schools in Norway were followed in the 2018–2019 academic year. Five of these classes were selected from schools applying traditional handwriting instruction from first grade (handwriting condition), and five of the classes were selected from schools using digital tablets in the writing instruction, postponing handwriting instruction (digital condition). The choice of basing the initial instruction on handwriting or digital tablets was anchored in policy at the district level. All schools were located in (semi-)rural parts in the west of Norway. (Semi-)rural areas in Norway are relatively socio-economically homogenous. National statistics indicated similar levels in family income and adult educational attainment. The sampled schools were state funded with equal or near-equal funding-per-pupil across schools and conditions. Class size across conditions was closely pair-matched.

Students were tested for literacy skills in their second to fifth weeks of school. Their ability to compose narratives was tested on five occasions, the first in November, and subsequently at intervals of approximately seven weeks, with the last assessment occasion in June. In the handwriting condition children wrote all texts by pencil and paper, while in the digital condition, students wrote all texts on keyboard on a digital tablet.

A total of 181 children (93 girls, 86 boys) participated in the study. At testing of literacy skills, a few students were absent, giving a final sample of 177 students (88 in the handwriting condition and 89 in the digital condition). Mean age at the time of testing of literacy skills (September) was 6 years, 3 months (*SD* = 3.4 months).

A few texts were missing or could not be analysed (e.g. not containing any words) on each of the five test occasions – ten, three, three, five, and seven texts, respectively.

2.2 Educational and instructional context

In Norway, children start school in August the year of their sixth birthday, and enrolment in first grade marks the onset of formal literacy instruction. The majority of Norwegian children, 97% of all five-year-olds, attend kindergarten (The Norwegian Directorate for Education and Training, 2019). No formal literacy instruction, for example, teaching of letters, is provided in kindergarten. Norwegian children are taught both handwriting and writing by keyboard in school. Schools are free to choose the order of handwriting and keyboarding instruction as long as children are taught to write in both modalities by the end of second grade. Therefore, both practices with and without digital devices co-exist in Norwegian first-grade classrooms, yielding an opportunity to examine how different practices affect children's learning to write.

Schools that were selected for the digital condition belonged to regions where the educational authority required that schools would provide all students with personal digital devices to be used in the initial instruction. Schools that were selected for the handwriting condition were schools that taught children with traditional methods involving pencil and paper, and that did not equip students with personal digital devices. All teachers answered a questionnaire the first and second semester concerning their instructional practice. Teachers in the handwriting condition reported that the writing instruction was based on pencil and paper: Learning letters and writing words, sentences, and texts involved writing by hand, while digital writing – in computer rooms or using shared sets of tablets/computers – only made up a minor part of their writing instruction. Teachers in the digital condition reported that writing-related activities typically occurred by digital tablet with little or no use of handwriting. The digital schools used different writing applications, but all used applications that allowed for text processing with text-to-speech synthesis. Teachers reported that they used default settings for the text-to-speech, which means that while writing the text-to-speech pronounced sounds, words, and sentences corresponding to what the students wrote. Also, all the written text could be read on demand. Other writing support offered by the applications, like spell-checker, was turned off during instruction and assessment tasks.

2.3 Literacy-ability measures

Students were tested individually by trained research assistants in a quiet room close to the students' classroom, and testing sessions lasted for approximately 20 minutes. The measures were extracted from a battery of tests previously used for Norwegian first graders (Lundetræ et al., 2017; Solheim et al., 2017, 2018). The spelling test was completed with pencil and paper, while a digital tablet was used for the rest of the literacy tests.

Grapheme-to-phoneme mapping

Twenty-four letters were presented one at a time in random order, and the children were asked to give the associated letter sound (Sunde et al., 2019). One point was given for each correctly sounded letter.

Phoneme isolation

The children were asked to identify the first sound of ten common Norwegian one- or two-syllable words (Haaland et al., 2023; Solheim et al., 2018). Pictures of the words were illustrated on the tablet screen, and the researcher named the objects, one at a time, and asked the child to say the first sound of the word. Each correctly isolated phoneme was awarded one point.

Phoneme blending

Children were instructed to blend a series of phonemes into a word (Solheim et al., 2018). Four images were shown on the tablet. A pre-recorded voice named all the objects (e.g. lys 'light', by 'city', fly 'airplane', sky 'cloud'). The pre-recorded voice then segmented one of the words (e.g. /f/, /l/, /y/) and the child was asked to point to the corresponding picture. The test consisted of eight test trials of increasing difficulty. Each correct answer gave one point.

Word reading

The students were asked to read aloud single words that appeared on the tablet screen (Haaland et al., 2023; Solheim et al., 2018). There were ten words of one or two syllables with regular spelling presented with increasing difficulty. Each correctly read word gave one point.

Spelling

The children were instructed to write single words from dictation with pencil on paper (Haaland et al., 2023; Solheim et al., 2018). There were ten tasks of increasing difficulty, from two-letter words to five-letter words. The researcher read a sentence and repeated the word that the student should write. All words were common Norwegian words with regular spelling. Words were scored as correct or incorrect.

Vocabulary

The ability to name objects was assessed through a short version of the Norwegian vocabulary test (Størksen et al., 2013) previously used in research (Solheim et al., 2018). Pictures of 20 items were presented one at a time on the tablet screen, and the students were asked to name the objects. Each correct answer was awarded one point.

2.4 Written composition assessment

Before the first writing task, children were introduced to the narrative genre. To make the writing motivating and authentic, children were introduced to a teddy bear, to whom they sent their stories, and throughout the year the bear responded with letters and pictures of him reading the stories. The writing tasks were researcher-designed, and each writing task consisted of a colourful picture, showing humans and/or animals in situations where something was about to happen (cf. Appendix A). Students were instructed to write a story based on the picture. They were also provided with three words representing central objects or actors in the picture (e.g. jente 'girl', tre 'tree', and ball 'ball'). Tasks were counterbalanced across test occasions, and task order was counterbalanced across conditions. Teachers administered the writing based on detailed instructions from the researchers sent by mail and e-mail prior to each test occasion. They were instructed to reserve one school lesson (45 minutes) for the writing to make sure that all students had enough time to complete their composition. The 45-minute lesson included a short introduction to the task. Students completing their composition earlier were given another task. Teachers were instructed to motivate the students to do their best, but not to help them with spelling or generating content for the stories.

Handwritten texts were transcribed into typescript. Students' spelling, spacing, and punctuation were kept, but reversed letters were corrected (except <b>/<d> and <p>/<q> substitutions, as we could not be sure these were reversals or spelling errors). In order to transcribe space between words, the space had to be bigger than the space between letters within the words. To transcribe space within a word (spacing error), there had to be enough space to insert a character.

All texts were blinded for condition and time and coded for nine text features (cf. below). These features reflect both lower-level skills like spelling and spacing, and higher-level skills like syntactic complexity and narrative features. At this age we would expect a high correlation between text length and higher-level features. Still, it would theoretically be possible to write a long text without much syntactic complexity or any narrative features. The three narrative measures were coded by the first author. In addition, 20% of the texts were double-coded by the second and the third author who coded 10% each. The sample of double-coded texts consisted of one text from each child, and task, time-point, and modality were equally represented.

Text length

Text length was measured as the total number of words written. Letter strings that, given context, represented two or more words were counted as separate words. Words that were incorrectly separated by spaces, given context, were treated as single words. Words containing spelling errors were included in the word count. Letter strings bounded by spaces that could not be identified as words were not included in the count.

Segmentation (spacing) accuracy

Spacing accuracy was operationalised as the number of correctly inserted spaces divided by the sum of correct, additional, and missing spaces. Correct spaces refer to correctly inserted spaces between words, additional spaces are incorrectly inserted spaces (within words), and missing spaces refer to locations where a space is required but missing. Punctuation, without spacing, was accepted as correct segmentation.

Sentence terminator accuracy

Terminator accuracy was operationalised as a count of correct terminators divided by the sum of correct, additional, and missing terminators. Correct terminators refer to correctly used sentence-terminating punctuation (full-stop, question mark, exclamation mark), additional terminators are terminators used incorrectly (e.g. full stop in the middle of a sentence), and missing terminators reflect locations where a sentence terminator was required but absent.

Spelling accuracy

Spelling accuracy was measured as the total number of correctly spelled words divided by the number of words written. Spelling errors did not include spacing errors.

Vocabulary sophistication

The measure of vocabulary sophistication was based on estimated age-of-acquisition ratings (e.g. Kuperman et al., 2012), but with a focus on written acquisition (Spilling et al., 2022). All content words (nouns, verbs, and adjectives) from the students' texts were extracted, lemmatised, and rated for written-age-of-acquisition. The 845 lemmas to be rated were divided into three online surveys and distributed to teachers and trainee teachers (N = 21, 16, and 16). For each word, the respondents were asked to indicate the age at which they believed the word would typically appear in children's written texts, on a scale from 5 to 14 years. The mean score across all the ratings for each word gave a written-age-of-acquisition score for that word. Each student's text then received a vocabulary-sophistication score based on the mean score of all content words (types) in the text. Internal reliability among raters for the three groups was .93 [.92, .94], .92 [.90, .93], and .94 [.92, .95], respectively (Cronbach's alpha, with 95% CI).

Syntactic accuracy and complexity (syntax)

All texts received a syntax score where every syntactically correct main clause yielded 1 point, every main clause with one or more syntactic errors gave 0.5 points, every syntactically correct subordinate clause gave 2 points, and every subordinate clause with one or more syntactic errors was credited 1.5 points. The syntax coding was done independently of the student's punctuation, and the main defining criterion for clauses was the presence of a verb.

Basic narrative structures (event count)

Narratives relate to a series of events (Labov & Waletzky, 1967), and the number of events was, therefore, used as a measure of basic narrative structures. All clauses denoting something happening at a point in time were coded as events. The first author coded all texts, while the second and third author doubled-coded 20% of the texts (10% each, in total 177 texts), with very good inter-rater reliability (intra-class correlation = 1, 95% CI [.99, 1]).

Advanced narrative structures

The content in stories can be linked through more advanced features than time, for example, thorough causal relations. Inspired by Martin and Rose (2008), these advanced narrative structures were identified: problem, solution, reaction, effect, and comment from the narrator. For every text, the number of each of the narrative structures was counted and then summed as a measure of the ability to use advanced narrative features. Coding of this feature was performed as the coding of event count, with very good inter-rater reliability (intra-class correlation = .97, 95% CI [.96, .98]).

Story grammar

A typical narrative is organised through several structural components centred around a complicated action (Labov & Waletzky, 1967). As the texts to be analysed were written by beginning writers, the most basic stages of story grammar – introduction, complication, and resolution – comprised the measure of story grammar. Coding of this feature was performed as the two other narrative measures with intra-class correlation = .95, 95% CI [.94, .97].

3. Statistical Analysis

Bayesian multivariate mixed-effects models (Gelman et al., 2014; Lambert, 2018) were used to determine whether the rate of change in composition performance depends on a combination of literacy skills and the modality – handwriting or typing – in which students were taught to write. The various outcome measures were therefore modelled jointly while giving separate parameter estimates for each measure. Because observations were nested within participants, and participants were nested within schools, all models with multiple time points were fitted with random intercepts for schools and students nested within schools, and with random by-school and by-student slope adjustments for the effect of time (intra-class correlations are given in Appendix B). All count data – number of words written, syntactic complexity, event count, and advanced narrative structures – were modelled as negative-binomial distributions to handle overdispersion. Ordinal data, for example story grammar, were modelled as cumulative distribution (Bürkner & Vuorre, 2019), and a normal distribution was used for the vocabulary sophistication measure. Modelling of the three accuracy measures was done with binomial distributions. Models were fitted with weakly informative priors for all effects. Weakly informative priors were specified to exclude implausible parameter values given the expected distributions of the outcome measures, without incorporating substantive theoretical assumptions or evidence from previous studies (Lambert, 2018).

Modelling used the probabilistic programming language Stan (Stan Developmental Team, 2022), accessed via the R brms package (Bürkner, 2018). All models were run with 20,000 iterations on 3 chains with a warm-up of 10,000 iterations and no thinning. Model convergence was confirmed by the Rubin–Gelman statistics (Gelman & Rubin, 1992) and by visual inspection of the sampling paths of the Markov Monte Carlo chains. Model comparison was quantified by estimation of predictive accuracy using leave-one-out cross validation (Vehtari et al., 2017). The models' predictive accuracy was compared using the difference in their expected log-predictive densities ($\Delta\widehat{(elpd)}$). This technique penalises the number of parameters to prevent overfitting. Parameter estimates are reported with their associated 95% probability intervals (the interval that contains the true parameter value with a 95% probability, cf. Sorensen et al., 2016).

Bayes factors (BF; e.g. Dienes, 2014; Wagenmakers et al., 2018) were calculated to quantify the evidence for the null hypothesis over the alternative hypothesis (BF01) and for the alternative hypothesis against the null hypothesis (BF10). A BF01 of 3, for example, indicates that the null hypothesis is three times more likely than the alternative hypothesis given the data. As a rule of thumb, a BF of 3 is interpreted as weak evidence, a BF of 5 as moderate evidence, and a BF of 10 as strong evidence (e.g. Jeffreys, 1961; Lee & Wagenmakers, 2014). BFs were calculated using the Savage–Dickey method (Dickey & Lientz, 1970).

4. Results

We first established the effects of literacy-skill measures on text composition at the first test occasion to determine the simple effect of literacy skills ignoring any effects of subsequent instruction. Findings from this analysis were then used to select measures to carry forward into the main analysis. We wanted to focus on measures that on the first time point could be predicted by literacy skills to build as informative models as possible when exploring effects on learning. In the main analysis we explored the effects of literacy skills on learning rate, and further whether these effects were moderated by the modality in which children were instructed and tested. This was achieved through comparison of a series of nested models, starting with an intercept-only model and then across seven further models adding effects of time (test occasion from 1 to 5), literacy measures, and modality, and their two and three-way interactions.

4.1 Effects of literacy skills on text composition at first test occasion

Observed means and correlations among the literacy measures are provided in Table 1. Grand means for all text-composition measures, and their residual correlations can be found in Table 2. We estimated the effects of literacy skills on text composition at the first test occasion by comparing an intercept-only model (with random school intercepts) and a model that included all literacy predictors. The latter model showed a higher predictive performance ($\Delta\widehat{elpd}$ = -

92.0, *SE* = 2.7) indicating that, overall, there is a predictive relationship between literacy measures and text composition. All literacy effects for all text-composition measures are shown in Appendix C. We carried forward into the analysis of learning over time just those literacy measures (grapheme–phoneme mapping and vocabulary) and text measures (text length, spelling accuracy, syntax, event count, and advanced narrative structures) that showed evidence for effects at the first test occasion.

When evaluating whether the two literacy predictors alone could predict text composition (the five selected text measures) at the first test occasion, we compared a zero model to a model that included the two literacy predictors. Adding literacy skills as predictors improved the model fit ($\Delta\widehat{elpd}$ = -67, *SE* = 20).

Table 1. Literacy skills measures: means and bivariate correlations

	Mean (SD)	First sound segmentation	Blending	Spelling	Vocabulary	Word reading
First sound segmentation	6.2 (3.6)					
Blending	3.6 (2.6)	.55				
Spelling	2.2 (3.1)	.60	.57			
Vocabulary	13 (3.7)	.43	.40	.48		
Word reading	3.4 (3.4)	.65	.58	.84	.49	
Grapheme–phoneme mapping	11 (6.8)	.71	.58	.82	.47	.84

4.2 Interaction effects of literacy skills and modality on learning composition

Based on the initial analysis we selected grapheme–phoneme mapping and vocabulary as the literacy predictors and text length, spelling accuracy, syntax, event count and advanced narrative structures as outcome measures for the main analysis. We built eight nested models allowing for main effects and interaction effects of time, literacy and modality, and compared models stepwise (listed in Table 3).

The complete model, Model 7, captures the hypothesis that we want to test, namely that start-of-school literacy affects the rate of learning text composition differently depending on the modality in which children are taught to write.

Results showed that the model that included time increased the predictive performance compared to the zero model (Model 1 vs Model 0 in Table 3). Adding the literacy skills improved the model fit (Model 2 vs. Model 1 in Table 3), and adding the interaction of time and literacy skills further improved the model fit (Model 3 vs Model 2 in Table 3). Adding modality as a main effect (Model 4 vs Model 3 in Table 3) gave better fit, but the difference in predictive performance was very small. After controlling for two-way interactions (Model 5 and Model 6) we found that adding three-way time-by-literacy skill-by-modality interactions (Model 7) did not improve model fit.

Parameter estimates from the complete model (Model 7) can be found in Table 4. Figure 1 illustrates estimated development in text composition over time for literacy skills and writing by hand and by keyboard. For visualisation we divided the literacy skill into three groups (children with low, average and high abilities). Our main interest was to explore whether the effect of modality on learning text composition (if any) is dependent on students' literacy skills. We found no evidence that different modalities affected students' rate of learning differently depending on their start-of-school literacy skills. Data showed evidence in favour of no effect, with strong evidence in support of no effect for seven out of the ten possible effects, and moderate or weak evidence in support of no effect for the other three possible effects (Table 4).

Further, looking at the other effects from the complete model (Table 4 and Figure 1), we found strong evidence in favour of a main effect of time on all text measures except for spelling accuracy. This suggests that there was a tendency for children to improve across the school year. Mean performance across the year tended to be greater for children who performed better on the grapheme–phoneme-mapping task at the start of the year. This was true for all text measures. Mean performance across the year also tended to be greater for children scoring higher on the vocabulary task at school start for the measures syntax and text length. Data did not, however, provide evidence that students' literacy skills affected the rate at which they learn to compose text across the year – Bayes factors showed moderate or strong evidence in favour of no interaction effect between time and literacy skills for nine out of ten possible effects, and inconclusive evidence for the last possible effect. We found that, on average, digital texts were better spelled than handwritten texts. That is, there was a small main effect of modality on spelling accuracy, with BF = 2.4 in favour of a non-zero effect. Further, we found that for mean performance across the year children experienced no (dis)advantage writing in either modality depending on their level of literacy skills: There was moderate or strong evidence in favour of no effect for nine out of ten possible effects, and inconclusive evidence for the interaction effect between modality and vocabulary on spelling accuracy. Finally, modality did not affect students' learning of text composition – Bayes factors showed moderate to strong evidence against there being an effect. All the findings presented in this section are consistent with our previous analysis of these data in Spilling et al. (2023).

Table 2. Estimated grand means and residual correlation among the outcome measures at the first test occasion. 95% probability intervals in brackets.

Text composition measure	Estimated mean	Terminator accuracy	Spacing accuracy	Spelling accuracy	Vocabulary soph.	Story grammar	No. of words	Syntax	Events
Advanced structures	1.42 [0.9, 2.28]	-.02 [-.09, .06]	-.09 [-.17, -.02]	-.24 [-.31, -.16]	.26 [.20, .32]	.70 [.66, .73]	.61 [.57, .66]	.59 [.55, .64]	.70 [.66, .73]
Events	1.61 [1.02, 2.45]	-.02 [-.10, .05]	-.14 [-.22, -.06]	-.19 [-.26, -.11]	.27 [.20, .33]	.61 [.56, .65]	.72 [.68, .75]	.72 [.68, .75]	
Syntax	2.11 [1.44, 2.97]	.03 [-.06, .11]	-.14 [-.22, -.05]	-.19 [-.27, -.11]	.33 [.26, .39]	.51 [.45, .56]	.90 [.89, .92]		
No. of words	13.43 [9.97, 18.55]	-.02 [-.11, .06]	-.16 [-.24, -.07]	-.20 [-.28, -.13]	.35 [.28, .41]	.53 [.48, .58]			
Story grammar	0.72 [0.51, 0.99]	.00 [-.07, .08]	-.08 [-.15, -.00]	-.16 [-.24, -.09]	.17 [.10, .24]				
Vocabulary soph.	6.86 [6.77, 6.95]	-.03 [-.11, .04]	-.12 [-.19, -.04]	-.32 [-.39, -.25]					
Spelling accuracy	0.71 [0.66, 0.77]	.10 [.02, .17]	.28 [.21, .35]						
Spacing accuracy	0.94 [0.89, 0.97]	.07 [-.01, .14]							
Terminator Accuracy	0.09 [0.03, 0.25]								

Table 3. Model specifications and model comparison

Model	Parameter added	Comparison	$\Delta\widehat{elpd}$ (SE)
Model 0	[intercept-only]		
Model 1	time	Model 1 vs Model 0	-2.9 (5.8)
Model 2	literacy skills	Model 2 vs Model 1	-1.2 (6.0)
Model 3	time : literacy skills	Model 3 vs Model 2	-11.6 (7.7)
Model 4	modality	Model 4 vs Model 3	-0.6 (2.6)
Model 5	literacy skills : modality	Model 5 vs Model 4	3.8 (3.5)
Model 6	time : modality	Model 6 vs Model 5	3.3 (2.6)
Model 7	time : literacy skills : modality	Model 7 vs Model 6	1.9 (4.4)

Note. All models were fitted with random intercepts for classrooms and students nested within classrooms, and with random by-class and by-student slope adjustments for the effect of time. Negative differences in the differences in estimated log-predictive density ($\Delta\widehat{elpd}$) show an improvement of the more complex model. Positive differences show that the more complex model did not improve model fit. Differences are shown with standard errors (SE). Colons ':' indicate interactions.

Table 4. Parameter estimates (Est.) with 95% probability intervals (PI) from Model 7 with statistical support in favour of no effect (BF01) and in favour of an effect (BF10) indicated by Bayes Factors (BF), where † marks moderate to strong evidence. Colons ':' indicate interactions.

	Advanced structures			Events			Syntax			No. of words			Spelling accuracy		
Predictor	Est. [95% PI]	BF01	BF10	Est., [95% PI]	BF01	BF10	Est., [95% PI]	BF01	BF10	Est., [95% PI]	BF01	BF10	Est., [95% PI]	BF01	BF10
Time	0.28 [0.13, 0.41]	< 0.1	13.7†	0.27 [0.15, 0.39]	< 0.1	53.4†	0.38 [0.27, 0.49]	< 0.1	> 100†	0.29 [0.19, 0.39]	< 0.1	> 100†	0.1 [-0.04, 0.24]	4.8	0.2
Modality	-0.09 [-0.48, 0.32]	5.2†	0.2	-0.02 [-0.47, 0.42]	5.3†	0.2	0.15 [-0.24, 0.53]	4.2	0.2	0.19 [-0.15, 0.54]	3.2	0.3	0.38 [0.05, 0.7]	0.4	2.4
Graph–phon map.	0.17 [0.06, 0.27]	0.2	4.4	0.15 [0.02, 0.27]	1.2	0.8	0.27 [0.14, 0.4]	< 0.1	> 100†	0.17 [0.06, 0.28]	0.2	6.1†	0.16 [0.05, 0.27]	0.3	3
Vocabulary	0.11 [0, 0.23]	3	0.3	0.12 [0, 0.26]	2.9	0.3	0.17 [0.03, 0.31]	0.8	1.2	0.13 [0.02, 0.24]	1.3	0.8	0.05 [-0.07, 0.16]	13.7†	< 0.1
Time : modality	0.08 [-0.11, 0.27]	7.9†	0.1	0.09 [-0.08, 0.25]	6.9†	0.1	-0.01 [-0.15, 0.14]	17.1†	< 0.1	0 [-0.14, 0.14]	17.9†	< 0.1	-0.1 [-0.3, 0.1]	5.8†	0.2
Time : graph– phon map.	-0.03 [-0.12, 0.05]	18.7†	< 0.1	-0.01 [-0.08, 0.07]	28.1†	< 0.1	0 [-0.07, 0.07]	30.2†	< 0.1	0.02 [-0.04, 0.08]	30.8†	< 0.1	0.04 [-0.02, 0.11]	13.9†	< 0.1
Time : vocabulary	-0.04 [-0.13, 0.05]	17.4†	< 0.1	-0.04 [-0.12, 0.05]	18.9†	< 0.1	-0.09 [-0.17, -0.01]	2.2	0.4	-0.05 [-0.11, 0.01]	10.3†	0.1	0.06 [-0.01, 0.14]	7.4†	0.1
Modality : graph–phon map.	0.08 [-0.07, 0.24]	8.1†	0.1	0.12 [-0.05, 0.3]	5†	0.2	0.12 [-0.07, 0.3]	5.3†	0.2	0.11 [-0.04, 0.26]	5.6†	0.2	-0.08 [-0.23, 0.07]	8.5†	0.1
Modality : vocabulary	-0.01 [-0.18, 0.15]	13.6†	< 0.1	-0.06 [-0.24, 0.12]	9.8†	0.1	-0.04 [-0.24, 0.15]	10.2†	0.1	-0.07 [-0.22, 0.09]	9.8†	0.1	-0.17 [-0.33, -0.01]	1.6	0.6
Time : modality : graph–phon map.	-0.06 [-0.17, 0.06]	11.1†	< 0.1	-0.08 [-0.19, 0.02]	6†	0.2	-0.09 [-0.19, 0.01]	4.6	0.2	-0.05 [-0.13, 0.03]	13.4†	< 0.1	0.08 [-0.01, 0.17]	6.1†	0.2
Time : modality : vocabulary	-0.02 [-0.15, 0.11]	16†	< 0.1	0 [-0.12, 0.12]	18.4†	< 0.1	0.04 [-0.07, 0.15]	14.4†	< 0.1	0.04 [-0.04, 0.13]	16.3†	< 0.1	-0.05 [-0.15, 0.06]	14.7†	< 0.1

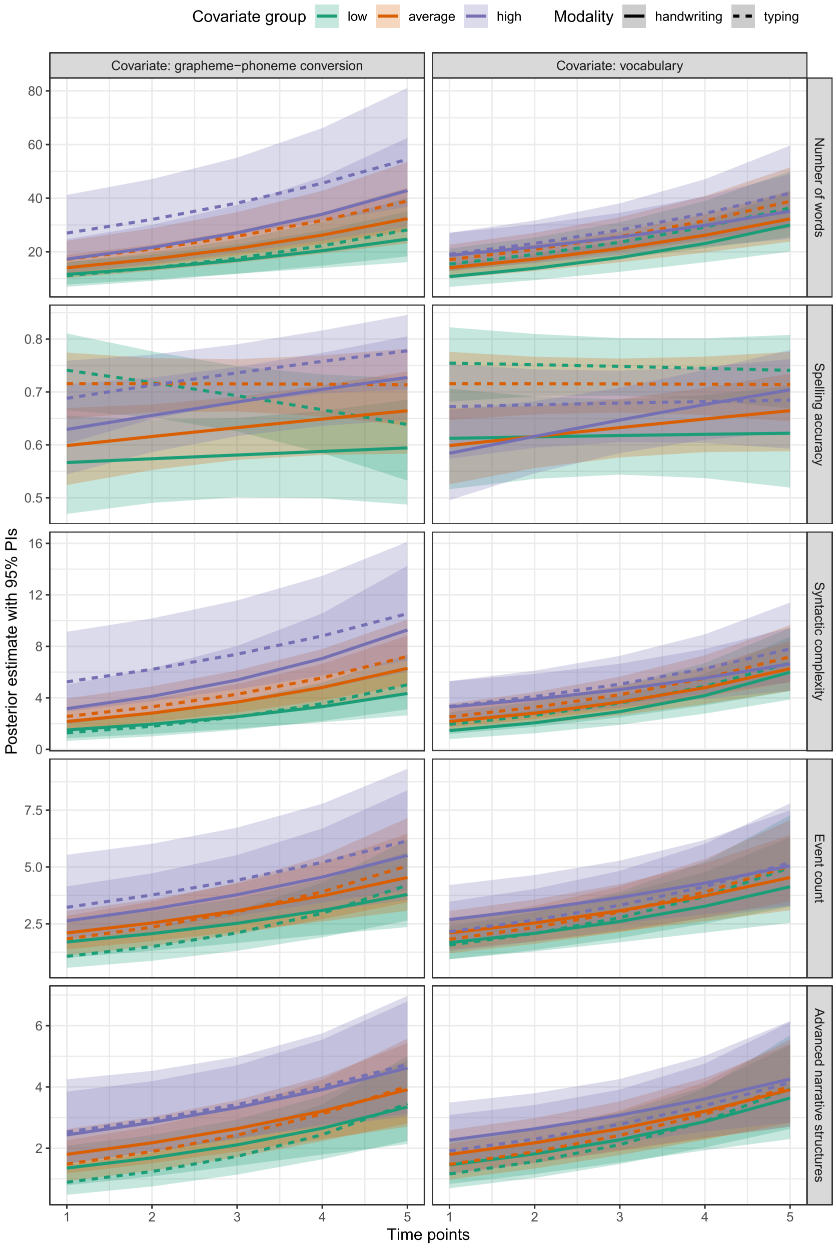

Figure 1. Learning rate estimates with 95% probability intervals (PIs). Outcome measures are shown in the row strips. Literacy skills are shown in the column strips.

5. Discussion

In the present study we explored whether the effect of modality on the rate of learning written composition is moderated by children's literacy skills at the start of school. Is it the case, for example, that learning to write by keyboard particularly benefits children with lower initial literacy skills?

Results showed that grapheme–phoneme-mapping ability and vocabulary size had (weak) effects on several text-composition measures both for performance at the first test occasion and for average performance across the school year. Observing correlations between literacy skills and composition performance is not unexpected given that such relationships have been found in much of the literature cited in the introduction (e.g., Kim et al., 2014). However, we found evidence that students' literacy skills do not influence the rate at which their composition skill developed across that year. This finding was somewhat surprising, as learning the skills necessary for producing a written narrative is a highly demanding task. It could, therefore, be the case that children who have come far in their development of literacy skills would benefit from this when starting school – unless they already know what is taught. For example, Torrance et al. (2021) found that spelling ability at school entry affected Spanish first-grade students' learning of written composition, where students with poor spelling ability showed initially slower but then more steady growth than their peers. However, in Spain, formal literacy instruction starts in kindergarten, and most students in the Torrance et al. (2021) study could already write simple words by the start of school. In the present study, the spelling ability of the participants probably resembled that of the minority of the Spanish students, who not until the first months of school developed sufficient spelling ability to produce multiple-sentence compositions.

Our main focus, however, was on whether modality and literacy-skill effect interacted over time. Note that this is possibly the case even given the failure to find overall effects of literacy skill on composition development. If, for example, a particular literacy skill acted to accelerate learning in one modality, but decelerate learning in the other modality, then these effects would cancel each other out, resulting in no overall effect of that literacy skill on learning writing. However, we found moderate to strong evidence against there being an interaction effect between modality and literacy skills on learning written composition. This suggests that when first graders learn to compose text, children with lower – or higher – literacy skills do not particularly benefit from learning to write by using pencil and paper or by typing on a digital device.

We hypothesised that children with various levels of literacy skills would benefit differently from the support provided by the digital tablet. For example, it could be that children with low grapheme–phoneme-mapping ability would benefit from writing by keyboard with visual letter representations as this might reduce cognitive demands. In contrast to our expectations, we did not find that these children produced more or higher-quality text when writing digitally. One explanation might be that for these novice children searching for and choosing the right letter among many other distracting letters when writing digitally may be both cognitively demanding and time consuming.

Another possible scenario could be that children with lower literacy skills were particularly aided in their spelling by the text-to-speech synthesis of the digital tablet, while children with higher literacy skills would show better performance over time when writing by hand. We did, however, not observe that children with different levels of literacy skills benefitted differently from the text-to-speech synthesis for spelling more correctly. One possible reason could be that although students had different levels of literacy skills, all were novice writers, and their literacy skills were generally at a very low level. Therefore, it seems that the support in spelling provided by the digital device was equally beneficial for all students.

We thought that particularly children who struggle with spelling would get more feedback on content in their texts when they wrote digitally, because they would make fewer errors. We do not have any knowledge of the exact feedback the children got, but we did not see more improvement on any text measure for students with lower literacy skills in the digital condition.

One reason for not finding a three-way interaction effect could be that our measures were not ideal for capturing this interaction. However, we argue that this was not the case. First, we found correlations between the predictors and composition performance. Second, the text measures reflected improvement over time. A more plausible explanation could be that the teacher differentiated the instruction well regardless of modality, resulting in similar rates of learning for all students. We cannot rule out the possibility that modality might interact with children's literacy skills and have an effect on text composition once children start developing more complex texts, or if students wrote under more demanding conditions, for example writing to a time constraint.

5.1 Limitations on generalisability and future research

The results of the present study must be understood in light of the particular context in which the study was conducted. Norwegian is a semi-shallow orthography, and children start school without any prior formal writing instruction. In contexts where children start school with more developed literacy skills findings might be different. Furthermore, our findings can only be generalised to digital contexts that are similar to those described in this study. Digital tools are in a state of flux, with AI being the major shift since this study was conducted (Anson & Cole, 2026). However, to date, writing instruction in first grade is not much affected by AI.

The present study did not include measures of students' motor skills. One argument for postponing handwriting and learning to write by keyboard is that it relieves the burden of learning to form letters at the same time as learning the cognitive processes involved in writing, which can be especially advantageous for students with slow motor development (Genlott & Grönlund, 2013). Future research could, therefore, investigate if students with poor graphomotor skills would benefit from learning to write by typing.

Also, studies investigating the further course of children's writing development when writing in different modalities would be advisable. As children develop as writers, they will be more proficient in transcription, and their writing will to a higher extent be constrained by ideation skills and higher-order cognitive skills, for example, reasoning (Juel et al., 1986; Kim

& Park, 2019). At the same time, children are expected to produce more sophisticated compositions, and they will have to continue developing their writing skills, for example, their skills in revising text on a substantial level. Future research could examine if – when students have automatised their writing skills – there is any advantage of well-developed vocabulary skills for students writing by keyboard.

5.2 Conclusion

Our study suggests that effects of writing modality (pencil and paper or keyboarding) on students' learning of written composition across first grade is not dependent on the literacy skills with which they start the school year. For example, it is not the case that children starting the year with poor letter knowledge particularly benefit from learning by keyboarding (or by pencil and paper). It is possible that these findings do not generalise to all educational contexts. Similar studies in other educational contexts are therefore necessary. This caveat aside, our findings are not consistent with claims that keyboarding (or handwriting) scaffolds learning to write for children with weaker general literacy ability by, for example, reducing cognitive load. For teachers considering whether to move away from traditional pencil-and-paper-based first-grade writing instruction towards increased use of keyboarding, our findings suggest that this decision will neither benefit nor disadvantage children with weaker literacy skills at this age.

Author note

This study was funded by the Research Council of Norway, Grant number 273422. We want to thank all participating children, teachers and schools. A manual for transcription and coding of the texts, data, and analysis scripts are available at Open Science Foundation: https://osf.io/9jud6/overview

Editor's note

One of the authors of this paper is also an editor of the Journal of Writing Research. To avoid potential bias in the review process, both the action editor and the reviewers were blind to this author's identity.

References

Albuquerque, A., & Alves Martins, M. (2022). Invented spelling as a tool to develop early literacy: The predictive effect on reading and spelling acquisition in Portuguese. *Journal of Writing Research*, *14*(1), 113–131. https://doi.org/10.17239/jowr-2022.14.01.04

Anson, C., & Cole, K. (2026). Empirical studies of writing and generative AI: Introduction to the special issue. *Journal of Writing Research*, *17*(3). https://doi.org/10.17239/jowr-2026.17.03.01

Babayiğit, S., & Stainthorp, R. (2011). Modeling the Relationships Between Cognitive-Linguistic Skills and Literacy Skills: New Insights From a Transparent Orthography. *Journal of Educational Psychology, 103(1),* 169–189. https://doi.org/10.1037/a0021671

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious Mixed Models*. http://arxiv.org/abs/1506.04967

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., Abbott, R. D., Augsburger, A., & Garcia, N. (2009). Comparison of pen and keyboard transcription modes in children with and without learning disabilities. *Learning Disability Quarterly, 32(3),* 123–141. https://doi.org/10.2307/27740364

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.

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal, 10(1)*, 395–411. https://doi.org/10.32614/RJ-2018-017

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal Regression Models in Psychology: A Tutorial. *Advances in Methods and Practices in Psychological Science, 2(1),* 77–101. https://doi.org/10.1177/2515245918823199

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(3),* 471–488. https://doi.org/10.1037/0022-0663.98.3.471

Dickey, J. M., & Lientz, B. P. (1970). The Weighted Likelihood Ratio, Sharp Hypotheses About Chances, the Order of a Markov Chain. *The Annals of Mathematical Statistics, 41(1),* 214–226.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5(781),* 1–17. https://doi.org/10.3389/fpsyg.2014.00781

Duiser, I. H. F., Ledebt, A., van der Kamp, J., & Savelsbergh, G. J. P. (2022). Does learning to write and type make a difference in letter recognition and discrimination in primary school children? *Journal of Cognitive Psychology, 34(6),* 691–702. https://doi.org/10.1080/20445911.2022.2060240

Dunsmuir, S., & Blatchford, P. (2004). Predictors of writing competence in four to seven year old children. *British Journal of Educational Psychology, 74(3),* 461–483. http://discovery.ucl.ac.uk/46215/

Fitjar, C. L., Rønneberg, V., Nottbusch, G. & Torrance, M.. (2021). Learning handwriting: Factors affecting pen-movement fluency in beginning writers. *Frontiers in Psychology, 12,* Article 663829. https://doi.org/10.3389/fpsyg.2021.663829

Fiskerstrand, P., & Gamlem, S. M. (2024). Mapping oral feedback interactions in young pupils' writing. *Assessment in Education: Principles, Policy & Practice, 31(3–4),* 204–220. https://doi.org/10.1080/0969594X.2024.2386515

Gamlem, S. M., Rogne, W. M., Rønneberg, V., & Uppstad, P. H. (2020) Study protocol: DigiHand – the emergence of handwriting skills in digital classrooms. *Nordic Journal of Literacy Research 6(2),* 25–41. https://doi.org/10.23865/njlr.v6.2115

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian Data Analysis* (3rd ed.). Chapman Hall/CRC.

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science, 7(4),* 457–72. DOI: 10.1214/ss/1177011136

Genlott, A. A., & Grönlund, Å. (2013). Improving literacy skills through learning reading by writing: The iWTR method presented and tested. *Computers & education, 67*, 98–104. https://doi.org/10.1016/j.compedu.2013.03.007

Genlott, A. A., & Grönlund, Å. (2016). Closing the gaps – Improving literacy and mathematics by ict-enhanced collaboration. *Computers & Education, 99*, 68–80. https://doi.org/10.1016/j.compedu.2016.04.004

Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002. *Journal of Technology, Learning, and Assessment, 2(1),* 1–52. https://ejournals.bc.edu/index.php/jtla/article/view/1661

Haaland, V. F., Rege, M., & Solheim, O. J. (2023). Do students learn more with an additional teacher in the classroom? Evidence from a field experiment. *The Economic Journal, 134(657)*, 418–435, https://doi.org/10.1093/ej/uead074

Islam, M. S., & Grönlund, Å. (2016). An international literature review of 1:1 computing in schools. *Journal of Educational Change, 17(2),* 191–222. https://doi.org/10.1007/s10833-016-9271-y

Jeffreys, H. (1961*). The Theory of Probability (Vol. 3).* Oxford University Press, Clarendon Press.

Jiménez, J. E., & Hernández-Cabrera, J. A. (2018). Transcription skills and written composition in Spanish beginning writers: pen and keyboard modes. *Reading and Writing, 32(7).* https://doi.org/10.1007/s11145-018-9928-4

Jones, I., & Pellegrini, A. D. (1996). The effects of social relationships, writing media, and microgenetic development on first-grade students' written narratives*. American Educational Research Journal, 33(3)*, 691–718. https://doi.org/10.3102/00028312033003691

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(4),* 243–255. https://doi.org/10.1037/0022-0663.78.4.243

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

Kiefer, M., Schuler, S., Mayer, C., Trumpp, N. M., Hille, K., & Sachse, S. (2015). Handwriting or Typewriting? The Influence of Pen- or Keyboard-Based Writing Training on Reading and Writing Performance in Preschool Children. *Advances in Cognitive Psychology, 11(4),* 136–146. https://doi.org/10.5709/acp-0178-7

Kim, Y. S. G., Al Otaiba, S., Sidler, J. F., & 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., Otaiba, S. Al, 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. 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(5),* 1319–1343. https://doi.org/10.1007/s11145-018-9913-y

Kruschke, J. K. (2015). Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan (2nd ed.). Elsevier.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44(4),* 978–990. https://doi.org/10.3758/s13428-012-0210-4

Labov, W., & Waletzky, J. (1967). Narrative Analysis: Oral Versions of Personal Experience. In J. Helm (Ed.), *Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society.* American Ethnological Society.

Lambert, B. (2018). *A student's guide to Bayesian statistics.* Sage.

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Longcamp, M., Zerbato-Poudou, M. T., & Velay, J. L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica, 119(1),* 67–79. https://doi.org/10.1016/j.actpsy.2004.10.019

Lucero, M., Fernández, M. J., & Montanero, M. (2018). Teachers' written feedback comments on narrative texts in Elementary and Secondary Education*. Studies in Educational Evaluation, 59*, 158–167. https://doi.org/10.1016/j.stueduc.2018.07.002

Lundetræ, K., Solheim, O. J., Schwippert, K., & Uppstad, P. H. (2017). Protocol: 'On track', a group-randomized controlled trial of an early reading intervention. *International Journal of Educational Research, 86,* 87–95. https://doi.org/10.1016/j.ijer.2017.08.011

Lunsford, A. A., & Lunsford, K. J. (2008). "Mistakes Are a Fact of Life": A National Comparative Study. *College Composition and Communication, 59(4),* 781–806. https://www.jstor.org/stable/20457033

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(7–8),* 643–672. https://doi.org/10.1023/A:1012071514719

Malpique, A. A., Valcan, D., Pino-Pasternak, D., Ledger, S., Asil, M., & Teo, T. (2023a). The keys of keyboard-based writing: Student and classroom-level predictors of keyboard-based writing in early primary. *Contemporary Educational Psychology, 75,* Article 102227. https://doi.org/10.1016/j.cedpsych.2023.102227

Malpique, A. A., Valcan, D., Pino-Pasternak, D., Ledger, S., & Kelso-Marsh, B. (2023b). Shaping young children's handwriting and keyboarding performance: Individual and contextual-level factors. *Issues in Educational Research, 33(4)*. http://www.iier.org.au/iier33/malpique.pdf

Malpique, A. A., Valcan, D., Pino-Pasternak, D., Ledger, S., & Merga, M. (2023c). Effect sizes of writing modality on K-6 students' writing and reading performance: a meta-analysis. *Australian Educational Researcher, 51(5),* 2001–2030. https://doi.org/10.1007/s13384-023-00676-y

Martin, J., & Rose, D. (2008). Genre Relations. *Mapping Culture.* Equinox.

Mayer, C., Wallner, S., Budde-Spengler, N., Braunert, S., Arndt, P. A., & Kiefer, M. (2020). Literacy Training of Kindergarten Children With Pencil, Keyboard or Tablet Stylus: The Influence of the Writing Tool on Reading and Writing Performance at the Letter and Word Level. *Frontiers in Psychology, 10*, Article 3054. https://doi.org/10.3389/fpsyg.2019.03054

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas – Part II. *Language and Linguistics Compass, 10(11),* 591–613. https://doi.org/10.1111/lnc3.12207

Oddsdóttir, R., Ragnarsdóttir, H., & Skúlason, S. (2021). The effect of transcription skills, text generation, and self-regulation on Icelandic children's text writing. *Reading and Writing, 34(2),* 391–416. https://doi.org/10.1007/s11145-020-10074-w

Poole, D., M. & Preciado, M. K. (2016). Touch typing instruction: Elementary teachers' beliefs and practices. *Computers & Education, 102*, 1–14. https://doi.org/10.1016/j.compedu.2016.06.008

Puranik, C., Lonigan, C., & Kim, Y.-S. (2011). Contributions of emergent literacy skills to name writing, letter writing, and spelling in preschool children. *Early Childhood Research Quarterly, 26(4),* 465–474. https://doi.org/10.1016/j.ecresq.2011.03.002

Read, J. C. (2007). A study of the usability of handwriting recognition for text entry by children. *Interacting with Computers, 19(1*), 57–69. https://doi.org/10.1016/j.intcom.2006.08.009

Rijlaarsdam, G., & Couzijn, M. (2000). Writing and learning to write: A double challenge. In R. J. Simons (Ed.), *New learning* (pp. 157–190). Kluwer.

Rogne, W. M., Rønneberg, V., Gamlem, S. M., Spilling, E. F., & Uppstad, P. H. (2024). Effects of digitalisation on learning to write – A naturalistic experiment. *Learning and Instruction, 93*, Article 101970. https://doi.org/10.1016/j.learninstruc.2024.101970

Rønneberg, V., & Nilsen, C. (2022). Teachers' talk about giving feedback to young text writers, and about giving feedback on handwritten and typed texts. *Writing and Pedagogy, 13(1–3),* 207–226. https://doi.org/10.1558/wap.21501

Rønneberg, V., Torrance, M., Uppstad, P. H., & Johansson, C. (2022). The process-disruption hypothesis: how spelling and typing skill affects written composition process and product. *Psychological research, 86(7),* 2239–2255. https://doi.org/10.1007/s00426-021-01625-z

Solheim, O. J., Frijters, J. C., Lundetræ, K., & Uppstad, P. H. (2018). Effectiveness of an early reading intervention in a semi-transparent orthography: A group randomised controlled trial. *Learning and Instruction, 58*, 65–79. https://doi.org/10.1016/j.learninstruc.2018.05.004

Solheim, O. J., Rege, M., & McTigue, E. (2017). Study protocol: "Two teachers". A randomized controlled trial investigating individual and complementary effects of teacher-student ratio in literacy instruction and professional development for teachers. *International Journal of Educational Research, 86,* 122–130. https://doi.org/10.1016/j.ijer.2017.09.002

Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *The Quantitative Methods for Psychology, 12(3),* 175–200. https://doi.org/10.20982/tqmp.12.3.p175

Spilling, E. F., Rønneberg, V., Rogne, W. M., Roeser, J., & Torrance, M. (2022). Handwriting versus keyboarding: Does writing modality affect quality of narratives written by beginning writers? *Reading and Writing, 35*, 129–153. https://doi.org/10.1007/s11145-021-10169-y

Spilling, E. F., Rønneberg, V., Rogne, W. M., Roeser, J., & Torrance, M. (2023). Writing by hand or digitally in first grade: Effects on rate of learning to compose text. *Computers & Education, 198*, Article 104755. https://doi.org/10.1016/j.compedu.2023.104755

Stan Developmental Team. (2022). *Stan Modeling Language Users Guide and Reference Manual, 2.30*. https://mc-stan.org.

Størksen, I., Ellingsen, I., Tvedt, M. S., & Idsø, E. M. C. (2013). Norsk vokabulartest (NVT) for barn i overgangen mellom barnehage og skole: Psykometrisk vurdering av en nettbrettbasert test [Norwegian Vocabulary Test (NVT) for children in the transition between kindergarten and school: Psychometric evaluation of a tablet-based test]. *Spesialpedagogikk Forskningsdel, 4(13)* 40–54.

Sunde, K., Furnes, B., & Lundetræ, K. (2019). Does introducing the letters faster boost the development of children's letter knowledge, word reading and spelling in the first year of school? *Scientific Studies of Reading, 24(2)*, 141–158. https://doi.org/10.1080/10888438.2019.1615491

The Norwegian Directorate for Education and Training. (2019). *The Norwegian Education Mirror 2019: Kindergartens.* https://www.udir.no/in-english/education-mirror-2019/school/#children-in-kindergarten

The Norwegian Directorate for Education and Training (2022). *The Norwegian Education Mirror 2022: Digital infrastructure and the school working day.* https://www.udir.no/in-english/the-education-mirror-2022/the-digital-state-of-schools-and-kindergartens/digital-infrastructure-and-the-school-working-day/

Torrance, M., Arrimada, M., & Gardner, S. (2021). Child-level factors affecting rate of learning to write in first grade. *British Journal of Educational Psychology, 91(2),* 714–734. https://doi.org/10.1111/bjep.12390

van Galen, G. P. (1991). Handwriting: Issues for a psychomotor theory*. Human Movement Science, 10(2–3)*, 165–191. https://doi.org/10.1016/0167-9457(91)90003-G

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27(5),* 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical Advantages and Practical Ramifications. *Psychonomic Bulletin & Review, 25(1),* 35–57. https://doi.org/10.3758/s13423-017-1343-3

Whitehurst, G., & Lonigan, C. J. (1998). Child Development and Emergent Literacy. *Child Development, 69(3),* 848–872. https://doi.org/10.1111/j.1467-8624.1998.tb06247.x

Appendix A: Writing tasks

Children were given a copy of this or a similar picture and were asked to write a story. In addition to the picture, students were provided with three central words, e.g.: jente 'girl', ball 'ball', and tre 'tree'. These words were placed below the picture.

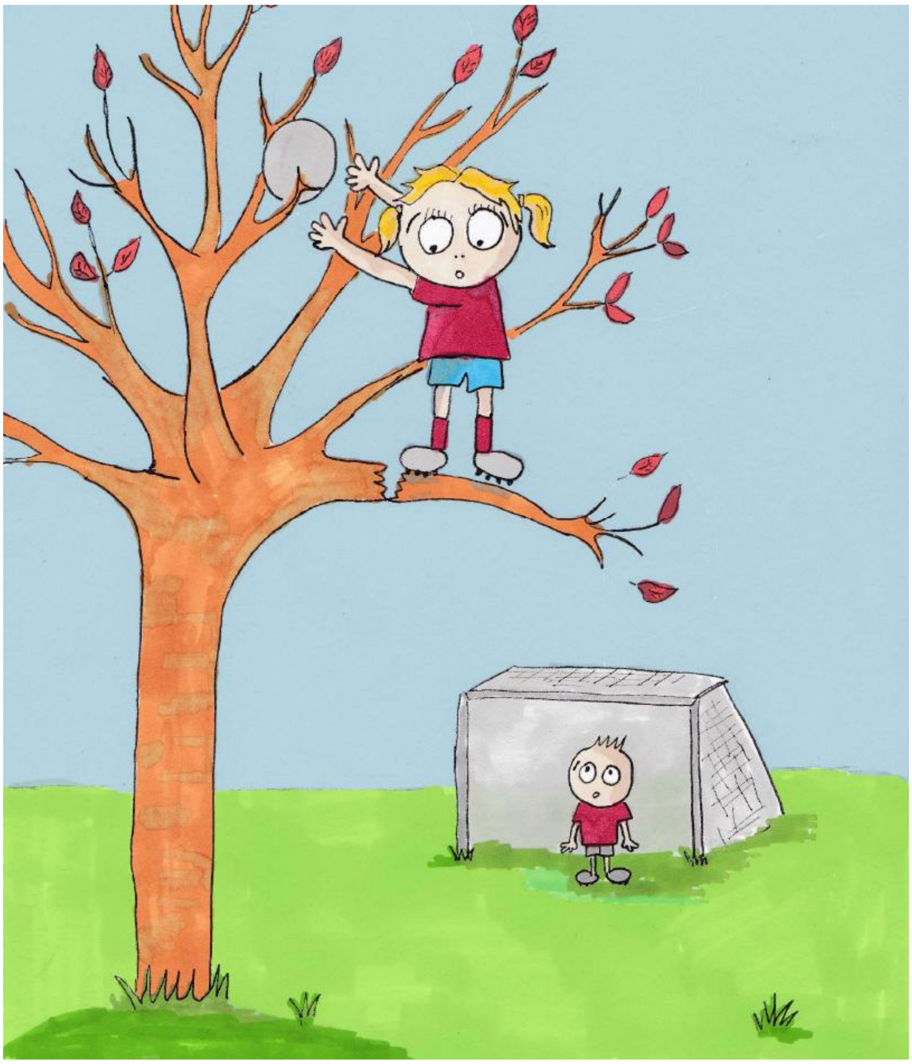

Figure A2. Example of writing task

Appendix B: Intra-class correlations (ICC) for children and schools by written composition measure

Table B1. Intra-class correlations for children

Composition measure	ICC children for all varying intercepts and their random time slopes
Advanced structures	0.035
Advanced structures : time	0.009
Events	0.056
Events : time	0.009
Syntax	0.067
Syntax : time	0.009
Number of words	0.049
Number of words : time	0.008
Spelling accuracy	0.043

Table B2. Intra-class correlations for schools

Composition measure	ICC for schools
Advanced structures	0.026
Events	0.033
Syntax	0.023
Number of words	0.019
Spelling accuracy	0.014

Appendix C: Effects of all literacy skills on all text measures at the first test occasion

Table C1. Effects of all literacy skills on all text measures at the first test occasion, estimates and 95% PI

	Advanced Structures	Events	Syntax	No. of words	Spelling accuracy	Story grammar	Vocabulary soph.	Spacing accuracy	Terminator accuracy
Graph–phon map.	0.13 [-0.18, 0.43]	0.18 [-0.12, 0.48]	0.3 [0.03, 0.58]	0.17 [-0.03, 0.38]	-0.24 [-0.41, -0.07]	-0.04 [-0.73, 0.64]	-0.04 [-0.17, 0.09]	0.37 [0.13, 0.62]	0.19 [-0.28, 0.66]
First sound seg.	0.07 [-0.15, 0.29]	0.2 [-0.01, 0.42]	0.17 [-0.04, 0.37]	0.09 [-0.05, 0.24]	-0.23 [-0.36, -0.1]	0.38 [-0.1, 0.87]	0.11 [0.02, 0.2]	-0.37 [-0.55, -0.19]	-0.03 [-0.34, 0.27]
Blending	0.06 [-0.12, 0.24]	0.05 [-0.13, 0.22]	0.06 [-0.11, 0.24]	-0.02 [-0.15, 0.1]	0.01 [-0.1, 0.12]	-0.05 [-0.44, 0.34]	0.05 [-0.02, 0.13]	-0.14 [-0.28, 0]	-0.32 [-0.62, -0.03]
Word reading	0.1 [-0.18, 0.37]	0.09 [-0.18, 0.37]	0.09 [-0.17, 0.34]	0.06 [-0.13, 0.26]	0.26 [0.1, 0.41]	0.53 [-0.11, 1.18]	0.1 [-0.03, 0.23]	0.29 [0.08, 0.51]	0.23 [-0.22, 0.68]
Spelling	0 [-0.24, 0.25]	-0.17 [-0.42, 0.07]	-0.1 [-0.34, 0.13]	-0.04 [-0.23, 0.15]	0.15 [0, 0.29]	0.24 [-0.35, 0.83]	-0.12 [-0.24, 0]	-0.06 [-0.25, 0.14]	-0.27 [-0.71, 0.16]
Vocabulary	0.21 [0.03, 0.39]	0.21 [0.04, 0.39]	0.22 [0.05, 0.38]	0.13 [0.01, 0.24]	-0.11 [-0.22, -0.01]	0.41 [0.01, 0.82]	0.05 [-0.02, 0.12]	-0.19 [-0.34, -0.04]	-0.15 [-0.4, 0.1]