CrossMark

Cognitive predictors of shallow-orthography spelling speed and accuracy in 6th grade children

Vibeke Rønneberg[1,2] · Mark Torrance[2,3]

Abstarct Spelling accuracy and time course was investigated in a sample of 100 Norwegian 6th grade students completing a standardized spelling-to-dictation task. Students responded by keyboard with accurate recordings of response-onset latency (RT) and inter-keypress interval (IKI). We determined effects of a number of child-level cognitive ability factors, and of word-level factors—particularly the location within the word of a spelling challenge (e.g., letter doubling), if present. Spelling accuracy was predicted by word reading (word split) performance, non-word spelling accuracy, keyboard key-finding speed and short-term memory span. Word reading performance predicted accuracy just for words with spelling challenges. For correctly spelled words, RT was predicted by non-word spelling response time and by speed on a key-finding task, and mean IKI by non-verbal cognitive ability, word reading, non-word spelling response time, and key-finding speed. Compared to words with no challenge, mean IKI was shorter for words with an initial challenge and longer for words with a mid-word challenge. These findings suggest that spelling is not fully planned when typing commences, a hypothesis that is confirmed by the fact that IKI immediately before within word challenges were reliably longer than elsewhere within the same word. Taken together our findings imply that routine classroom spelling tests better capture student competence if they focus not only on accuracy but also on production time course.

✉ Vibeke Rønneberg
vibeke.ronneberg@uis.no

Mark Torrance
mark.torrance@ntu.ac.uk

1   University of Bergen, Bergen, Norway

2   University of Stavanger, Stavanger, Norway

3   Nottingham Trent University, Nottingham, UK

🖄 Springer

Introduction

Both practitioners and researchers typically understand spelling competence as ability to spell words accurately. Spelling success is, however, only partly about whether or not the word that the child produces is accurately spelled. Students also need to spell fluently, without excessive hesitation and effort. Factors that affect children's spelling fluency—the time course of single-word production—are not well understood. Understanding the full complexity of a student's spelling ability arguably requires information about both the product and production time course.

There are both educational and theoretical reasons for developing an understanding of factors that predict spelling fluency. A child who struggles in an attempt to spell all words correctly is probably more disadvantaged when composing text than a child who writes all words quickly, but makes a few mistakes. At a global level, struggling with spelling production may result in students being demotivated, running out of time, having less time for planning or writing a shorter text. At a local level, difficulty with specific words tends to slow output. For example, copy-typists tend to slow down when they type irregular words (Bloemsaat, Van Galen, & Meulenbroek, 2003). Struggling to spell a word—regardless of whether the word is then spelled correctly—risks damage to higher level, conceptual, rhetorical or even syntactic structures in the text. Language processing operates under a stringent "now-or-never" constraint (Christiansen & Chater, 2016). If spelling is attention demanding, and therefore draws processing resources away from higher-level processes, the writer may, in a literal sense, forget what they were going to say. This trade-off between transcription (spelling and handwriting) and higher-level conceptual or rhetorical processing has frequently been argued (e.g., Berninger, 1999; McCutchen, 1996; Torrance & Galbraith, 2006; von Koss Torkildsen, Morken, Helland, & Helland, 2016), although evidence of a causal relationship between spelling competence and text quality (spelling accuracy aside) is not yet established (see, for example, Babayiğit & Stainthorp, 2010; Graham & Santangelo, 2014).

From a theoretical perspective, exploring spelling time course addresses issues around the nature and scope of the planning processes necessary for written production of a single word. In principle at least, spelling can be achieved either by assembly—incremental phoneme-letter mapping—or by activation of orthographic lexemes directly from their associated concepts (e.g., Rapp, Benzing, & Caramazza, 1997). This suggests a dual-route account of orthographic retrieval (Barry, 1994; Martin & Barry, 2012; Perry & Ziegler, 2004; Rapp, Epstein, & Tainturier, 2002). The extent to which different routes are used (or the probability that a particular route will win the horse-race, cf. Paap & Noel, 1991) is likely to be language-dependent. Share (2008) argues in the context of single-word reading that the dual-route model as a whole is not a good representation of reading in a shallow orthography. Similar arguments may apply to shallow-orthography spelling—the

focus of the present study. A related issue is the scope of lexical and motor planning in single word production. Specifically, is all orthographic processing complete prior to output (before the writer makes their first pen stroke or key stroke; e.g., Logan & Crump, 2011) or does orthographic planning persist beyond output onset? Again, it is possible that the answer to this question is dependent on the orthographic depth of the language that is being written. In developing writers both processing route (assembly vs. direct) and planning scope (lexical or sublexical) are likely to be dependent on various dimensions of the writer's literacy skills.

Both orthographic and phonological processing ability are likely, therefore, to play a role in children's spelling performance. A number of studies have explored cognitive factors that predict early spelling performance, where performance is measured just on the basis of accuracy. In early primary school children there is, as might be expected, a strong correlation between single-word reading accuracy and spelling accuracy. This effect is present in both shallow and deep orthographies (Babayiğit & Stainthorp, 2011; Caravolas, Hulme, & Snowling, 2001) but in German it disappears in early-secondary students (Landerl & Wimmer, 2008). This is because the regularity of letter-to-sound mapping in German mean that in older students, word reading errors are very rare, while spelling mistakes remain relatively common as sound-letter mapping is less regular. Spelling accuracy is also consistently predicted by phonological awareness (e.g., Muter, Hulme, Snowling, & Stevenson, 2004) although these effects also decrease with age in typically-developing students (Furnes & Samuelsson, 2011; Landerl & Wimmer, 2000). In English, rapid automatized naming (RAN)—the number of stimuli (letters, numbers, colours, objects) that a participant can name in a fixed period of time—has also been found to predict spelling accuracy, even after control for phonological ability (Savage, Pillay, & Melidona, 2008; Strattman & Hodson, 2005), although Landerl and Wimmer (2008) failed to find similar effects in German in 8th grade students. While RAN may play a role in predicting spelling accuracy, the shared mechanisms are not well understood. It may be that RAN taps speed of phonological processing (Vellutino, Fletcher, Snowling, & Scanlon, 2004). Alternatively, slow naming may indicate impaired learning of mental orthographic representations, at either word or sub-word level (Bowers & Newby-Clark, 2002; Wimmer, Mayringer, & Landerl, 2000; but see Moll, Fussenegger, Willburger, & Landerl, 2009). Moll et al. (2014) found that RAN measures involving naming digits and objects (i.e. with no letter-to-phoneme processing component) predicted spelling accuracy in English but not French (deep orthographies), and German but not Finnish (relatively transparent orthographies, although with some asymmetry in German, noted above). These phonological and orthographic theories make different predictions about the extent to which RAN will predict spelling accuracy and/or spelling speed for regular and irregular words (i.e. words with and without predictable letter-sound correspondence) and, more generally, spelling in shallow and deep orthographies. However McGeown, Johnston, and Moxon (2014) found that even in English where spelling is highly irregular, spelling accuracy in mid-primary children is predicted both by pseudo-word reading—a test of phonological decoding—and by orthographic ability (based on a word/pseudo homophone discrimination task). Finally, there is some evidence that verbal short-term memory

span predicts spelling accuracy for early-primary spelling in Norwegian (Lervåg & Hulme, 2010), and early- and mid-primary children spelling in English (Caravolas et al., 2001), although it is not clear whether these effects are independent of phonemic awareness (Landerl & Wimmer, 2008).

The present study investigated effects of various child-level and word-level factors on not just spelling accuracy but also production fluency. In general terms, this examined accuracy/speed trade-off in spelling-to-dictation: Are accuracy and fluency predicted similarly by students' orthographic and phonological skills, or do effects diverge? More specifically, analysis of how student skill interacts with word regularity provides insight into the cognitive processes that underlie upper-primary students' performance on spelling-to-dictation tasks. There is considerable evidence from research with adult writers that spelling processes and their motor execution interact in single word production (Bertram, Tønnessen, Strömqvist, Hyönä, & Niemi, 2015; Delattre, Bonin, & Barry, 2006; Kandel & Perret, 2015; Lambert, Kandel, Fayol, & Espéret, 2007; Roux, McKeeff, Grosjacques, Afonso, & Kandel, 2013; Scaltritti, Arfé, Torrance, & Peressotti, 2016). Delattre et al., for example, found that in French (a deep orthography), both word frequency and orthographic regularity affected onset latency (time from stimulus presentation to output-onset, measured as onset of first pen stroke) and output duration (time from onset to word end, controlling for word length). Kandel and Perret (2015) found similar effects in 8-year-old children, an age at which most are likely to have achieved handwriting automaticity. They concluded that orthographic processing persists beyond output onset. Scaltritti et al. (2016) in a study of adults typing picture names in Italian (an orthographically transparent language) found frequency effects and other lexical effects on response latency. They found, however, that orthographic neighbourhood —a word-level orthographic effect—did not affect onset latency but did affect production duration. Torrance et al. (2017) studied typed picture naming in a number of alphabetic languages. Spelling difficulty, indexed by cross-subject spelling agreement and controlling for name agreement, only affected writing timecourse in some languages, including French but not including Norwegian (the focus of the present study). In almost all cases, effects were exclusively in production duration and not onset latency. There is evidence that extended production time is specifically associated with slowed output around the location of the irregularities within the word (Roux et al., 2013).

In general terms, therefore, findings suggest (a) individual (student-level) differences in orthographic and/or phonological processing ability predict spelling accuracy in early primary students, and (b) orthography and orthography/phonology regularity affect processing timecourse—both response latency and output duration —in adults and also late-primary students. Correlational studies exploring student-level predictors of spelling performance however give quite mixed findings which may, in part, be explained by differences in orthographic depth of the languages in which the research was conducted. Most obviously, the role of word-level orthographic knowledge is likely to be greater in deep orthographies in which spelling by assembly is relatively unreliable. In a cross-language comparison of students beyond second grade using nationally standardised spelling tests (i.e. without attempt to control regularity across language), Moll et al. (2014) found that

phonological awareness predicted spelling accuracy in German and Finnish (relatively transparent orthographies), but not in English and French (deep orthographies), but also not in Hungarian, which is relatively transparent. It may also be that the orthography in which a student learns to spell affects the extent to which direct or assembled routes dominate their spelling processes: independently of the regularity of the specific word being spelt students with a shallow-orthography language may prefer spelling by assembly.

The study that we report in this paper explores the effects of Norwegian 6th grade students' performance on RAN (letter), phonological (non-word) spelling, and a word-split reading task that emphasises orthographic recognition, on their spelling performance, measured in terms of both accuracy and fluency. Students completed a spelling-to-dictation task and typed their responses. We also measured, as control variables, key-finding ability as an index of typing skill, and non-verbal ability and RAN (objects) as a measure of general cognitive skills and short-term memory. Norwegian, like German, is typically understood as somewhat asymmetrical in terms of orthographic transparency (e.g., Hagtvet & Lyster, 2003; Lervåg & Hulme, 2010). Grapheme-phoneme mapping is very regular, making reading by assembly accurate. Phoneme-grapheme correspondence is less predictable, meaning that for a subset of Norwegian words spelling by assembly alone is insufficient, and success also requires retrieval of orthographic knowledge. Spelling in the present study was assessed via an existing, standardised spelling task (Skaathun, 2013) comprising both regular words with a simple 1-to-1 phoneme-letter mapping, and words that contained some form or irregularity or complexity (henceforth "challenge") for which spelling would necessarily need to go beyond incremental letter-by-letter assembly.

Writing, unlike speech, has two main output modalities—handwriting and keyboarding (typing). There has been a tendency to for researchers to theorise these independently (e.g., Gentner, Larochelle, & Grudin, 1988; van Galen, 1991). We know of no direct comparison of spelling processes in the two modalities. Arguably, however, it is reasonable to assume that processes upstream of motor output that is required for the generation of spelling is very similar in both modalities. Our choice of typed output in the present study was therefore expedient rather than principled: Writing time course data are more easily collected by keyboard, and give more clearly demarcated character onset times. Although students in Norwegian schools, with a small number of exceptions, learn handwriting before typing, students are expected to also have a reasonable level of typing competence by upper primary.

Our research addressed questions about the effects of child-level and word-level factors on spelling accuracy and timecourse, and the interaction between the two. After statistical control for general cognitive skills (object RAN and non-verbal ability), we predicted effects for both phoneme-to-grapheme encoding ability (non-word spelling-to-dictation) and orthographic recognition. That is, in a more shallow orthography we expected both phoneme-grapheme encoding and orthographic recognition to be important when spelling. More specifically, we predicted divergent effects for regular words and words that contain a spelling challenge; with students with good orthographic ability performing particularly well when spelling challenging words. We determined effects on accuracy but also, for

correctly spelled words, effects on production fluency, and particularly the extent to which spelling processes persists beyond typing onset. If planning is prepared fully in advance of typing onset then both response latency (RT, time from stimulus presentation to typing onset) and production speed should be unaffected by whether the spelling challenge occurs at the start of the word or in the middle of the word. If orthographic planning persists beyond typing onset then we predicted differential effects of challenge location. We also predicted that, if orthography is prepared incrementally (on a letter-by-letter basis) then inter-keypress interval (IKI) immediately prior to a spelling challenge would be longer than elsewhere in the same word.

Methods

Participants

Whole classes of Norwegian 6th grade students from four different schools were invited to take part in a "writing week" during which they completed a battery of measures including all those reported in this paper. Data collection was between mid-February and mid-April, 2015. Students were excluded from the sample if they did not speak Norwegian at home, they had behavioural difficulties that prevented successful completion of tasks, and if they were absent for part of the test period (20 students in total). 100 students (61 females) with a mean age of 11 years and 10 months made up the final sample. Students varied in keyboarding skill but all used computers regularly as part of normal classroom activities, including free writing tasks. Mean within-word inter-key interval on a free-writing task completed by the same sample (not reported in this paper) was 380 ms. This compares to 167 ms for Norwegian upper secondary students (Torrance, Rønneberg, Johansson, & Uppstad, 2016).

Materials and procedures

All tests were given to the participants at their respective schools. Some tests were administered in a group setting, while others were given individually, as detailed below. Group tests were all completed under "examination conditions". The first author administered all group tests. The first author and a research assistant administered individual tests.

Spelling assessment

Participants completed an existing spelling-to-dictation test, standardized for students writing by hand, and consisting of 32 items designed to cover different features of Norwegian orthography. The test was designed to measure spelling ability across a wide age range and had previously been standardised with large samples of Norwegian children (Skaathun, 2007, 2013 see "Appendix"). In our study, students wrote on a keyboard. The test comprises a combination of words

with a straightforward phoneme-grapheme mapping (7 words), words with a word-initial challenge (8 words), and a larger group of words with mid-word challenges (17 words). Challenges included consonant doubling following a short vowel sound (e.g., *tatt*/taken), consonant clusters (*marsjerer*/marches) these have been found to be difficult to spell (Hagtvet, Helland, & Lyster, 2005), failure to differentiate similar phonemes (e.g., [ʃ], [ʂ], and [ç] in the word *kjole*/dress), and silent letters (the letter *g* in *gjort*/done). Word frequency (surface form) was calculated from the Norwegian Newspaper Corpus (NNC, 2013). Word frequencies varied between 966 and <1 per million, with a mean of 225. The spelling test was completed by small groups of students on individual computers. Words were first presented in a sentence and then the children were told which word to spell. (e.g., "I use a comb to style my hair. Write comb"). The target word could appear anywhere in the sentence, meaning that participants could not infer which word they were going to write until it was repeated. Target words and sentences were pre-recorded and presented to participants through headphones. There was no time limit to complete the spelling test, but participants were urged to start spelling the word as soon as possible after the word was presented. To move on to the next word participants pressed enter. For each trial we recorded *accuracy* (whether or not the word was spelled correctly) and fluency. Fluency was measured in terms of, *response-onset latency* (RT, the time from the onset of the target word (e.g., onset of *comb* in *Write comb*) to first keypress), and mean *inter-keypress interval* (MIKI, the mean of all IKIs during production of the word, excluding time prior to first keypress and after last keypress).

Nonverbal ability

Students completed Raven's Standard Progressive Matrices (Raven, 1981) as a measure of general nonverbal cognitive ability. This was group-administered.

Rapid automatized naming (object and letter RAN)

Students were given the letters, and objects subtests from the CTOPP (Wagner, Torgesen, & Rashotte, 1999). Students were presented with two printed pages of each with 36 randomly arranged objects or letters. They were asked to name these as quickly as possible while the researcher scored accuracy and recorded time-to-completion. Students completed this task individually. Score is number of seconds to name all 72 items.

Word-split

A Norwegian version of the word-split task (Jacobson, 2001; Miller-Guron, 1999) was given to the participants as a fluency-focused measure of decoding ability. The word split test consists of 73 word strings each containing four words without inter-word spacing. Participants divide these word chains into single words by drawing a line between words, aiming to complete as many as possible within 5 min. The test is scored according to how many word chains the participants successfully manage

to solve. In a shallow orthography this gives a better measure of reading performance in older children than what is provided by single-word reading (Wimmer, 1993), a single word reading task stands the risk of reaching a ceiling effect in an orthography where most words can be read through assembly. The word split task puts emphasis on fluency in addition to accuracy. Because the words are presented with no spacing between them, and because of the time pressure students need to be able to recognize whole words or parts of words that typically make up part of a word, word endings for example in order to get a good score. The word-split task is believed to involve orthographic decoding and to a smaller degree grapheme-phoneme conversions. This task was group-administered.

Short-term memory

Students performed a letter-span task, comprising 3 practice sets followed by 10 experimental sets varying in length from 2 to 6 letters. Letter names were presented through headphones at 500 ms intervals. At the end of each set students repeated aloud all letter names they could recall, in order. If a student recalled all items in a set correctly then they scored the number of items in the set. If they failed to recall any of the items in the set then they scored zero. Scores were then summed across sets, giving a maximum possible overall score of 40.

Key finding

As a measure of keyboard familiarity, and therefore a proxy measure of keyboarding skill, we determined how quickly and accurately students were able to find single keys on a computer keyboard in response to spoken letter-name prompts. Students heard a total of 14 letter names, for letters across a range of frequencies, each occurring twice, with order randomized. On hearing a letter name they were required to press the corresponding key, being as quick and as accurate as possible. Students completed the exercise in small groups, on individual computers with letter names played through headphones. We recorded both *accuracy* (the total number of correctly chosen keys, with a maximum of 28) and *speed*—time from stimulus onset to keypress.

Non-word spelling

To assess phoneme-grapheme encoding, students completed a non-word spelling test comprising 20 pseudo-words varying in length from three letters to ten letters (see "Appendix"). Pseudo-words were created by using phonologically plausible letter-combinations that were easy to pronounce in Norwegian. For the shortest words only one letter separated the non-word from a real word (e.g., *fyt–fyr*). For the longer words the non-words were further away from real words. These stimuli had not been used in other studies. Procedure and scoring were the same as for the real-word spelling task. All phonetically plausible spellings were accepted as correct responses.

Results

Means and inter-measure bivariate correlations for child-level variables are given in Table 1. Mean spelling accuracy, across students, was 70.5% ($M = 22.6$, $SD = 4.1$), placing students, on average, at just below the 40th centile relative to national standards for their age group. Note however, that the test was standardized for students with no urge to start writing words as quickly as they could, and with the possibility to go back and edit previous words.

In describing our findings, we will first explore effects on spelling accuracy and then effects on initial and within-word latencies for words that were spelled correctly. In each case, we first report analyses of effects of child-level predictors. We then explore effects due to the sub lexical features of the target word (specifically location of potential spelling challenge).

Analysis throughout was by incremental sequences of mixed effects regression models, evaluated using the R lme4 package (Bates, Maechler, Bolker, & Walker, 2013). Model fits were evaluated by χ^2 change tests. Statistical significance of parameter estimates was established by z-test for the binomial logistic models reported in the first section of our results and by t test, with Satterthwaite approximation for denominator degrees-of-freedom, for models with continuous predictors (all other models). All continuous predictor variables were standardized prior to analysis. All chronometric variables (measures based on response latencies or inter-keypress intervals), and target-word frequency and length were log-transformed to reduce skew. Response latencies and inter-key intervals were trimmed at 2.5 SD.

Table 1 Means and bivariate correlations (r) for child-level predictors

	M (SD)	1	2	3	4	5	6	7	8
1. RAN objects	59.9 (13.7)								
2. Nonverbal reasoning (Raven)	42.2 (5.5)	−.15							
3. Non-word spelling accuracy	13.6 (2.8)	−.15	.29						
4. Non-word spelling RT (mean, ms)	2400 (429)	.18	−.07	−.09					
5. Word reading (word split task)	38.5 (11.0)	−.28	.38	.32	−.31				
6. Key-finding accuracy (proportion)	.90 (.07)	−.05	.01	−.07	−.03	.06			
7. Key-finding speed (mean, ms)	1259 (326)	.29	−.18	−.33	.45	−.47	.02		
8. RAN letters	31.5 (7.0)	.65	−.07	−.16	.22	−.23	.03	.29	
9. Short –term memory	23.6 (7.9)	−.12	−.02	.27	−.03	.14	−.05	−.11	−.19

$p < .001$ for $|r| > .19$

Effects on spelling accuracy

Child-level predictors

To determine what cognitive factors predict spelling accuracy, we tested logistic mixed effects regression models with spelling accuracy (child spelled word correctly vs. child did not spell word correctly) as the binomial dependent variable. We started with a zero (intercept-only) model, with random by-subject intercepts. Model 1 added predictors related to non-verbal ability and processing speed (Raven and RAN objects). By adding these variables first, we made sure general cognitive skills were controlled for when including the other predictor variables into the model. Model 2 added orthographic decoding (word-split scores). Model 3 added measures of phonological encoding ability (non-word spelling accuracy and response time). Improved fit as a result of adding these variables therefore indicate contribution of phoneme-grapheme encoding skill over and above that captured by the word-split score. Model 4 added factors relating to single-letter recognition and production (RAN letters, and key-finding speed and accuracy). Finally, Model 5 added short-term memory scores (STM).

Model fits are detailed in Table 2 and parameter estimates from the final model are given in Table 3. Both word-split performance and non-word spelling accuracy predicted spelling accuracy, after controlling for general (non-verbal) ability. Students with greater STM span also tended to spell more accurately. The final model also gave some evidence of greater accuracy for students who were quicker at key finding.

Word-level predictors

We determined the effects of the nature of the spelling challenge imposed by words in the spelling test—word-initial, mid-word, or none—statistically controlling for the effects of word length and frequency, using models similar to those described in

Table 2 Model fits for regression models with just child-level predictors of response accuracy, RT and within-word keypress latencies

	Correct response	Response latency (RT)	Mean within-word keypress latency (MIKI)
Model 1: zero model plus RAN Objects, Nonverbal Reasoning	$\chi^2(2) = 15.3$, $p < .001$	$\chi^2(2) = 6.9$, $p = .03$	$\chi^2(2) = 1.5$, $p = .005$
Model 2: model 1 plus Word Reading (word split task)	$\chi^2(1) = 23.2$, $p < .001$	$\chi^2(1) = 6.4$, $p = .012$	$\chi^2(1) = 33.4$, $p < .001$
Model 3: model 2 plus Non-word spelling accuracy and RT	$\chi^2(2) = 11.0$, $p = .004$	$\chi^2(2) = 3.5$, $p < .001$	$\chi^2(2) = 21.0$, $p < .001$
Model 4: model 3 plus Key-finding accuracy and speed and RAN Letters	$\chi^2(3) = 4.8$, $p = .185$	$\chi^2(3) = 6.5$, $p = .09$	$\chi^2(3) = 17.2$, $p = .001$
Model 5: model 4 plus Short-term memory	$\chi^2(1) = 7.20$, $p = .008$	$\chi^2(1) = 2.2$, $p = .14$	$\chi^2(1) = .64$, $p = .425$

Values are χ^2 change tests for comparison with previous model

Table 3 Child-level predictors: regression coefficients from the final model (Model 5)

	Correct response (probability)	Response latency (RT, ms)	Mean within-word key-press latency (MIKI, ms)
RAN objects	−.01 (−.04, .01)	26 (−25, 83)	1 (−15, 19)
Nonverbal reasoning (Raven)	.00 (−.02, .02)	−15 (−55, 29)	32 (17, 50)***
Word reading (word split task)	.04 (.02, .05)***	−2 (−47, 49)	−31 (−42, −17)***
Non word spelling accuracy	.02 (.00, .03)*	−42 (−80, 1)	−1 (−14, 14)
Non word spelling RT (mean, ms)	.01 (−.01, .02)	102 (56, 153)***	23 (9, 41)**
Key-finding accuracy	.01 (−.01, .02)	4 (−35, 46)	6 (−7, 20)
Key-finding speed (mean, ms)	−.02 (−.05, .00)*	61 (13, 114)**	33 (15, 52)***
RAN letters	−.01 (−.03, .01)	−28 (−75, 23)	8 (−7, 26)
Short-term memory	.02 (.01, .04)**	−31 (−68, 10)	−5 (−17, 8)

Parameters represent estimated change in dependent variable resulting from an increase of 1 SD in the predictor. 95% CI in parenthesis

* $p < .05$, ** $p < .01$, *** $p < .001$ for H$_0$: parameter = 0

the previous section. We started with an intercept-only model with random by-subject intercepts, and random slopes for word length and for the challenge-type factor. Random slopes for frequency did not improve model fit, and frequency and length were strongly collinear, so frequency slopes were omitted. We then added fixed factors incrementally as follows: We first added length and frequency as control variables, then challenge type (no challenge, word-initial, mid-word). We then added main effects for all child-level factors, and finally incrementally added interactions between the challenge-type factor and the child-level (individual differences) factors that showed main effects in the child-level analysis.

As might be expected, after control for frequency and length [Model 1, χ^2 (2) = 77, $p < .001$] adding challenge type substantially improved prediction of accuracy [$\chi^2(2) > 100, p < .001$]. Parameter estimates suggested that items with a mid-word challenge were 7.6% less likely to be spelled correctly [95% CI (2.5, 15.1), z = −3.4, $p < .001$] and items with a word-initial challenge were 6.4% more likely to be spelled correctly [95% CI (3.1, 7.3), z = 3.0 $p = .003$]. We found some evidence of an interaction between challenge-type and performance on the word-split task. Adding this effect improved model fit [χ^2 (2) = 5.97, $p = .050$]. These suggest little or no positive benefit for spelling accuracy of word-split ability for items without a spelling challenge, and that benefit were greater for items with mid-word and significantly better for word-initial challenges (z = 2.4, $p = .015$ from comparison with the no-challenge slope). There were no other statistically significant interactions between challenge type and child-level factors.

Effects on spelling fluency

In this section, we explore the process of spelling production, testing models similar to those described in the previous section but with response-onset latency (RT) and mean inter-keypress intervals (MIKI) as dependent variables.

Child-level predictors

We tested linear mixed effects models, starting with a baseline model with random by-subject and by-item intercepts. For RT (but not for MIKI) this model also included a fixed factor representing the duration of the audio presentation of the stimulus (i.e. the word to be spelled). Because RT was timed from stimulus onset, stimulus duration necessarily had a substantial effect on RT, and so required statistical control.

Incremental model fits and parameter estimates from the full model can be found in Tables 2 and 3. RT was predicted by the students' non-word spelling RT, and by speed on the key-finding task. There was some evidence of a word-split effect when this measure was added (Model 2), but this effect was subsumed by non-word spelling RT when these were added (Model 3). There were no other effects on RT. By contrast, speed of production once typing has commenced (MIKI) was predicted by word-split performance. MIKI was also predicted by non-word spelling RT and key-finding speed. Unexpectedly, students who got a high score for non-verbal cognitive ability were slower in spelling words, after the initial key. This effect was present in the final model (Table 3).

Word-level predictors

We constructed models following the same design as in the analysis of word-level effects on accuracy, starting with baseline models as described in the previous section, and then adding length and frequency effects as control variables [χ^2 (2) = 14, $p < .001$ and χ^2 (2) = 61, $p < .001$ for RT and MIKI respectively]. Adding challenge-type to the model did not improve fit for RT ($\chi^2 < 1$) but did improve fit for MIKI [χ^2 (2) = 68, $p < .001$]. Items with an initial challenge were produced an estimated 21 ms per character more quickly after the initial keystroke, relative to non-challenge words [95% CI (−32, −7)], and words with a mid-word challenge 31 ms per character more slowly [95% CI (17, 46)]. We found no interaction between challenge type and any of the child-level factors.

We hypothesized that the effect of a mid-word challenge on MIKI resulted from students pausing for longer when they reached the spelling challenge. To test this we looked at individual inter-keypress intervals (IKIs) just for items with mid-word challenges, comparing IKIs immediately prior to a challenge with IKIs at other locations (excluding word-initial latency). We added this factor to a linear mixed effects model with the challenge-type as a random by-subject slope and with random by-item and by-subject intercepts. This gave significantly improved fit [χ^2 (1) = 6.6, $p = .010$]. IKI immediately prior to the challenge was slower by an estimated 110 ms [estimated means with 95% CIs: challenge-initial, 480 (406, 569);

other, 370 (337, 406)]. There was no evidence of an effect on the latency associated with the keypress one key before the challenge-initial key. Note that by comparing keypress latencies within the same word this analysis controls for word-level differences in length and frequency.

Accuracy and process

Finally, we determined whether RT and MIKI were dependent on whether or not the word was spelt correctly. Again, just words that were produced fluently (i.e. with no editing) were included. We started with baseline models that included by-item intercepts, by-subject intercepts and slopes for word-length and for a factor representing whether or not the item was correctly spelled and fixed effects for word length and frequency. Adding a fixed effect of correct spelling gave significantly improved model fit for RT [χ^2 (1) $=$ 7.7, $p =$ .005] with longer pauses prior to initiation for words that ended up being incorrectly spelt [estimated means: correctly spelled words, 1973 ms (1888, 2061); wrongly spelled words, 2084 ms (1967, 2209)]. We found no evidence of a similar effect on MIKI.

Discussion

The present study differs from previous research exploring single-word spelling (Babayiğit & Stainthorp, 2011; Caravolas et al., 2001; Furnes & Samuelsson, 2011; Moll et al., 2014; Nikolopoulos, Goulandris, Hulme, & Snowling, 2006) in three respects, each of which affect interpretation of findings. First, students were spelling in a shallow orthography. This is likely to make accurate spelling easier, in general, but also it may make spelling by assembly (letter-by-letter mapping of phonemes onto graphemes) a more reliable and therefore more practiced route than is the case in deeper orthographies. Second, in contrast to the majority of studies of spelling development participants were at an age and stage where spelling skills can be expected to have been largely mastered, particularly given the transparent orthography in which they were writing. Third, students were completing an existing, standardized writing task. This has some benefits in that the effects of various factors on spelling accuracy can be indexed against population norms. Using an existing test however had the disadvantage that item-selection was less carefully controlled than we would have wished. With this context in mind, our results point towards the following conclusions.

Looking first at effects on accuracy we found, in line with our predictions, effects of both encoding accuracy (non-word spelling) and orthographic recognition (word-split) performance. These findings are consistent with the argument that both lexical retrieval and assembly play a role in spelling-to-dictation performance (Rapp et al., 2002), at least for a shallow orthography. Non-word spelling relies mainly on direct phoneme-grapheme mapping, word-split performance is underpinned by an ability to rapidly access orthographic lexical knowledge, and therefore is likely to be associated with direct-route processing. This suggests that it will predict spelling performance particularly in the case where the assembly route is likely to be slow or

fail. We found some evidence that this was the case: When comparing words with and without a challenge, word-split ability predicted spelling accuracy only for words that contained a spelling challenge.

Accuracy was also predicted by key-finding response time, but not key-finding accuracy, and by short-term memory span. The failure to find key-finding accuracy effects is probably best explained by the inadequacy of single-key finding as an index of accuracy of typing in a word-production context. This measure did not, for example, correlate with non-word spelling accuracy, suggesting that it does not predict tendency to make motor errors ("typos") in normal typing. Key-finding speed, on the other hand, showed some evidence of an effect, even after controlling for several other response-time and vigilance measures, including non-word spelling RT. If rapid key-finding is associated with relatively automatized keyboard skills, then this suggests that keyboarding has the potential to draw attention away from processing spelling, to the detriment of spelling accuracy. It is also probable that keyboarding fluency is associated with more writing practice—students who type quickly write more—which in turn might lead to better learning of spelling: Students who write more are exposed to more spelling decisions and more spelling-related teacher feedback.

Short-term memory span predicted accuracy, even after statistical control for all other predictors. This therefore appears to be a robust effect, which cannot be explained simply in terms of, for example, sustained attention, which was tapped by several other tasks. We do not have a straightforward explanation for this effect, beyond observing that performance on short-term memory span tasks is itself likely to depend, at least in part, on ability to recognize common patterns in the presented stimuli (Jones & Macken, 2015). This may represent a mechanism that memory-span performance and spelling production have in common, particularly given that in the present study the span stimuli were letters.

Two additional points are worth emphasizing here. First, the design of the study —multiple measures and incremental model building—makes it possible to isolate fairly precisely the effects of specific cognitive abilities, even when separate measurement tasks necessarily require a combination of these. Our failure to find RAN effects, for example, (contra Babayiğit & Stainthorp, 2011; Furnes & Samuelsson, 2011) may be because in the present study the various RAN component skills were subsumed by other measures (encoding, decoding, short-term memory). Second, sizes of effect in this study may appear quite small: A standard deviation increase in word-split gave a predicted improvement in accuracy of 3%— around 1 test point. However if should be noted that for 6th grade students, who have relatively high spelling competence, national norms for this spelling test indicate a range of just 4 points between the 50th and 80th centiles.

Effects of child-level factors were not limited to accuracy but extended to the speed with which correctly-spelled words were generated. RT (time to initiate response) was predicted just by response latency on the non-word spelling and speed on the key-finding task. We found no other effects. Particularly, we found no effect for word-split performance, including no interaction with the spelling challenge factor. MIKI (speed after typing onset) by contrast was predicted quite strongly by word-split performance, with an estimated increase of 31 ms (relative to

a mean of 370 ms) for a 1 *SD* change in word-split score. This latter finding is consistent with our argument that good word-split performance is associated with increased tendency to spell by retrieval of orthographic lexemes, and are therefore an increased tendency to complete preparation of the word in advance of typing onset. Whether this theory predicts longer (or shorter) word-initial latencies is not clear, and depends on the relative time costs of responding to the auditory word stimulus by initiating grapheme-phoneme conversion and by retrieving the associated orthographic lexeme.

More generally, the fact that students' mean onset latency on the non-word spelling task predicted within word MIKI when spelling real words suggests both that assembly played an important role in students' spelling, and that assembly cascaded beyond typing onset. Direct confirmation of this came from two, related findings. First, words that contained a mid-word spelling challenge were produced reliably more slowly than words with either no challenge or a word-initial challenge. Second, IKIs immediately before a mid-word spelling challenge were reliably longer by an estimated 110 ms (about 30%) relative to elsewhere in the same word. These findings suggest that spelling challenges were addressed, at least some of the time and for some students, when they were met during output and not prior to typing onset.

We did not find any evidence that challenge-type affected word-initial latency. However, it would be wrong to conclude on this basis that initial latency was independent of spelling difficulty. Preparation times for wrongly spelled words were reliably longer, by an estimated 107 ms, after statistical control for length and frequency. At minimum, this suggest that, again for some students some of the time, features of the words' orthography that make it easy or difficult to spell affect word-initial planning. Interestingly accuracy did not appear to be related to post-onset typing speed. One possible explanation is that stored inaccurate spellings compete at point-of-retrieval with correct spellings. In extreme cases (e.g., spelling *necessary* in English for at least one of the present authors) this competition reaches consciousness and is experienced as uncertainty. In the context of a shallow orthography, phoneme-grapheme mappings are likely to be very well learned, and so similar competition does not occur during (assembled) output. Also interestingly, Torrance et al. (2016) found the opposite effect in spontaneous text production, with shorter word-initial latencies for wrongly spelled words, again after control for length and frequency. This might be related to the more explicit focus on accuracy in a spelling test. Note, however, that in both Torrance et al.—which involved a free writing task—and in the present study—because of the use of an existing standardised task developed originally just as a measure of spelling accuracy— length and frequency were no controlled across words that varied in spelling difficulty. Although we held spelling and frequency as covariates in our analyses, stronger conclusions would have been possible had we built this control into the design of the task.

Two final findings: As might be expected, key-finding speed predicted typing speed. Less predictably, students who performed well on Raven's Matrices, as a general measure of non-verbal ability, typed more slowly, to a non-trivial extent. We do not have a straightforward explanation for this effect.

In summary, therefore, our findings suggest two general conclusions: First good pseudo-word encoding skills and ability to recognise orthographic patterns above the letter level are both associated with more accurate and more rapid spelling. This remains true after control for general-ability factors, and measures of keyboarding competence. This finding is consistent with an account of spelling-to-dictation in a shallow orthography, by late-primary aged students, that involves a combination of both phoneme $\rightarrow$ grapheme assembly, that cascades beyond typing onset, and retrieval of orthographic lexemes (necessarily in advance of typing onset). Second, these factors affect not only spelling accuracy but also spelling fluency even when words were spelled correctly: Students with good encoding and single-word reading skills spell more quickly.

The implication of this second finding for production of extended, spontaneous text—the narratives and short essays that are typical of upper primary education—is not immediately clear. Several authors have argued that there is a causal relationship between the ease with which a child can spell and their ability to generate coherent and ideationally rich text (Sumner, Connelly, & Barnett, 2013; von Koss Torkildsen et al., 2016). However, as we indicated in our introduction, evidence in support of this claim is mixed. If correlation exists between text quality and spelling fluency then third-factor explanations are perhaps most parsimonious: It may simply be that having a well-stocked and easily-accessed orthographic lexicon independently results in good spelling performance and good text-composition performance. At minimum, the present study suggests that spelling fluency, in addition to spelling accuracy, varies across students, is predicted by both orthographic recognition and sound-to-letter processing skill, is worthy of future research, and possibly also should be a focus of teacher attention.

Appendix

See Tables 4 and 5.

Table 4 Spelling test items

Challenge	Word	English
None	*eplepai*	apple-pie
	kam	comb
	lokomotiv	locomotive
	pløyer	plough
	risper	scratch
	sauer	sheep
	trikset	trick
Word-initial	*gi*	give
	gjelder	apply
	gjerne	happily
	gjort	done
	hjem	home
	hvor	where
	kilometer	kilometre
	kjole	dress
Mid-word	*belagt*	coated
	diesel	diesel
	dusj	shower
	fingrene	fingers
	forkjølet	cold
	godt	good
	grovt	wholemeal
	kanskje	maybe
	kråkeskrik	crow scream
	landet	country
	marsjerer	marches
	pizza	pizza
	rundt	around
	sprinklene	louvers
	tatt	taken
	tillegg	addition
	viktig	important

Table 5 Non word test items

Word
pok
fyt
låp
Aret
dese
glyfe
narot
kjave
keggi
fusse
kjermu
saruli
otelsu
muletla
baska
kaveløy
larkt
pafrsen
drogmulsio
josebjador

References

Babayiğit, S., & Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: A longitudinal study. *Reading and Writing, 23*(5), 539–568. doi: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. *Journal of Educational Psychology, 103*(1), 169–189. doi:10.1037/a0021671.

Barry, C. (1994). Spelling routes (or roots or rutes). In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 27–49). Chichester: Wiley.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-5.

Berninger, V. W. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly, 22,* 99–112. doi:10.2307/1511269.

Bertram, R., Tønnessen, F. E., Strömqvist, S., Hyönä, J., & Niemi, P. (2015). Cascaded processing in written compound word production. *Frontiers in Human Neuroscience, 9*((April)), 207. doi:10.3389/fnhum.2015.00207.

Bloemsaat, G., Van Galen, G. P., & Meulenbroek, R. G. J. (2003). Lateralized effects of orthographical irregularity and auditory memory load on the kinematics of transcription typewriting. *Psychological Research, 67*(2), 123–133. doi:10.1007/s00426-002-0112-7.

Bowers, P. G., & Newby-Clark, E. (2002). The role of naming speed within a model of reading acquisition. *Reading and Writing, 15,* 109–126. doi:10.1023/A:1013820421199.

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language, 45*(4), 751–774. doi:10.1006/jmla.2000.2785.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences, 39,* e62. doi:10.1017/S0140525X1500031X.

Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32*(6), 1330–1340. doi:10.1037/0278-7393.32.6.1330.

Furnes, B., & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences, 21*(1), 85–95. doi:10.1016/j.lindif.2010.10.005.

Gentner, D. R., Larochelle, S., & Grudin, J. (1988). Lexical, sublexical, and peripheral effects in skilled typewriting. *Cognitive Psychology, 20*(4), 524–548. doi:10.1016/0010-0285(88)90015-1.

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703–1743. doi:10.1007/s11145-014-9517-0.

Hagtvet, B., Helland, T., & Lyster, S. (2005). Literacy acquisition in Norwegian. In R. Joshi & P. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 15–30). Mahaw, NJ: Erlbaum.

Hagtvet, B. E., & Lyster, S.-A. H. (2003). The spelling errors of Norwegian good and poor decoders: A developmental cross-linguistic perspective. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 181–207). Philadelphia: Whurr Publishers.

Jacobson, C. (2001). *Manual: LäsKedjor.* Stockholm: Psykologiförlaget.

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition, 144,* 1–13. doi:10.1016/j.cognition.2015.07.009.

Kandel, S., & Perret, C. (2015). How does the interaction between spelling and motor processes build up during writing acquisition? *Cognition, 136,* 325–336. doi:10.1016/j.cognition.2014.11.014.

Lambert, E., Kandel, S., Fayol, M., & Espéret, E. (2007). The effect of the number of syllables on handwriting production. *Reading and Writing, 21*(9), 859–883. doi:10.1007/s11145-007-9095-5.

Landerl, K., & Wimmer, H. (2000). Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied Psycholinguistics, 21*(2), 243–262. doi:10.1017/S0142716400002058.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology, 100*(1), 150–161. doi:10.1037/0022-0663.100.1.150.

Lervåg, A., & Hulme, C. (2010). Predicting the growth of early spelling skills: Are there heterogeneous developmental trajectories? *Scientific Studies of Reading, 14*(6), 485–513. doi:10.1080/10888431003623488.

Logan, G. D., & Crump, M. J. C. (2011). Hierarchical control of cognitive processes. In Ross B. H. (Ed.), *Psychology of learning and motivation* (1st ed., Vol. 54, pp. 1–27). Elsevier Inc. doi: 10.1016/B978-0-12-385527-5.00001-2.

Martin, D. H., & Barry, C. (2012). Writing nonsense: The interaction between lexical and sublexical knowledge in the priming of nonword spelling. *Psychonomic Bulletin & Review, 19*(4), 691–698. doi:10.3758/s13423-012-0261-7.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review.* doi:10.1007/BF01464076.

McGeown, S. P., Johnston, R. S., & Moxon, G. E. (2014). Towards an understanding of how children read and spell irregular words: The role of nonword and orthographic processing skills. *Journal of Research in Reading, 37*(1), 51–64. doi:10.1111/jrir.12007.

Miller-Guron, L. (1999). *Wordchains: A word reading test for all ages.* Windsor: NFER-Nelson.

Moll, K., Fussenegger, B., Willburger, E., & Landerl, K. (2009). RAN is not a measure of orthographic processing. Evidence from the asymmetric german orthography. *Scientific Studies of Reading, 13*(1), 1–25. doi:10.1080/10888430802631684.

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., et al. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction, 29,* 65–77. doi:10.1016/j.learninstruc.2013.09.003.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*(5), 665–681. doi:10.1037/0012-1649.40.5.665.

NNC. (2013). Norwegian Newspaper Corpus v0.9. Retrieved from http://avis.uib.no/.

Nikolopoulos, D., Goulandris, N., Hulme, C., & Snowling, M. J. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 94*(1), 1–17. doi:10.1016/j.jecp.2005.11.006.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53*(1), 13–24. doi:10.1007/BF00867328.

Perry, C., & Ziegler, J. C. (2004). Beyond the two-strategy model of skilled spelling: Effects of consistency, grain size, and orthographic redundancy. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 57*(2), 325–356.

Rapp, B., Benzing, L., & Caramazza, A. (1997). The autonomy of lexical orthography. *Cognitive Neuropsychology, 14*(1), 71–104. doi:10.1080/026432997381628.

Rapp, B., Epstein, C., & Tainturier, M.-J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive Neuropsychology, 19*(1), 1–29. doi:10.1080/0264329014300060.

Raven, J. (1981). *Manual of raven progressive matrices and mill hill vocabulary scales (Res. Suppl. No. 1)*. London: HK Lewis & Co., Ltd.

Roux, S., McKeeff, T. J., Grosjacques, G., Afonso, O., & Kandel, S. (2013). The interaction between central and peripheral processes in handwriting production. *Cognition, 127*(2), 235–241. doi:10.1016/j.cognition.2012.12.009.

Savage, R., Pillay, V., & Melidona, S. (2008). Rapid serial naming is a unique predictor of spelling in children. *Journal of Learning Disabilities, 41*(3), 235–250. doi:10.1177/0022219408315814.

Scaltritti, M., Arfé, B., Torrance, M., & Peressotti, F. (2016). Typing pictures: Linguistic processing cascades into finger movements. *Cognition, 156,* 16–29. doi:10.1016/j.cognition.2016.07.006.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134*(4), 584–615. doi:10.1037/0033-2909.134.4.584.

Skaathun, A. (2007). *Staveferdigheit. Ei undersøking av stavetileininga i norsk barneskole*. Stavanger: Universitetet i Stavanger.

Skaathun, A. (2013). *Lesesenterets staveprøve*. Stavanger: Universitetet i Stavanger, Lesesenteret.

Strattman, K., & Hodson, B. W. (2005). Variables that influence decoding and spelling in beginning readers. *Child Language Teaching and Therapy, 21*(2), 165–190. doi:10.1191/0265659005ct287oa.

Sumner, E., Connelly, V., & Barnett, A. L. (2013). Children with dyslexia are slow writers because they pause more often and not because they are slow at handwriting execution. *Reading and Writing, 26*(6), 991–1008. doi:10.1007/s11145-012-9403-6.

Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 67–82). New York: Guilford 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*(5), 375–388. doi:10.1080/10888438.2016.1205071.

Torrance, M., Nottbusch, G., Alves, R. A., Arfé, B., Chanquoy, L., Chukharev-Hudilainen, E., & Wengelin, Å. (2017). Timed written picture naming in 14 European languages. *Behavior Research Methods*. doi:10.3758/s13428-017-0902-x.

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

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 45*(1), 2–40. doi:10.1046/j.0021-9630.2003.00305.x.

von Koss Torkildsen, J., Morken, F., Helland, W. A., & Helland, T. (2016). The dynamics of narrative writing in primary grade children: Writing process factors predict story quality. *Reading and Writing, 29*(3), 529–554. doi:10.1007/s11145-015-9618-4.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *CTOPP: Comprehensive test of phonological processing*. Austin: Pro-ed.

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics, 14*(1), 1–33. doi:10.1017/S0142716400010122.

Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology, 92*(4), 668–680. doi:10.1037//0022-0663.92.4.668.