Adolescent weak decoders writing in a shallow orthography: Process and product

Journal:	*Scientific Studies of Reading*
Manuscript ID	HSSR-2015-0049.R3
Manuscript Type:	Research Article
Keywords:	dyslexia, fluency, text production, writing process, keystroke logging

Adolescent weak decoders writing in a shallow orthography: Process and product

Mark Torrance, Nottingham Trent University, UK and University of Stavanger, Norway; mark.torrance@ntu.ac.uk; Corresponding Author

Vibeke Rønneberg, University of Bergen and University of Stavanger, Norway: vibeke.ronneberg@uis.no

Christer Johansson, University of Bergen, Norway; christer.johansson@lle.uib.no

Per Henning Uppstad, University of Stavanger; per.h.uppstad@uis.no

It has been hypothesised that students with dyslexia struggle with writing because of a word-level focus that reduces attention to higher level textual features (structure, theme development). This may result from difficulties with spelling and/or from difficulties with reading. 26 Norwegian upper secondary students (*M* = 16.9 years) with weak decoding skills and 26 age-matched controls composed expository texts by keyboard under two conditions: normally and with letters masked to prevent them reading what they were writing. Weak decoders made more spelling errors and produced poorer quality text. Their inter key-press latencies were substantially longer pre-word, at word-end, and within-word. These findings provide some support for the word-level focus hypothesis, although we found that weak decoders were slightly *less* likely to engage in word-level editing. Masking did not affect differences between weak decoders and controls indicating that reduced fluency was associated with production rather than monitoring what they had produced.

Running Head: Weak Decoders' Writing Processes

Introduction

Decoding refers to the translation of letters and words on the page into mental lexical representations. Struggling with decoding is a central diagnostic characteristic of dyslexia (Lyon, Shaywitz, & Shaywitz, 2003), and the negative effects of decoding deficits on comprehension of multi-sentence text are relatively well established (e.g., Ransby & Swanson, 2003). Much less is known about how dyslexia in general, and weak decoding in particular, affects students' performance when asked to compose their own multi-sentence texts.

Students with dyslexia typically experience difficulty with producing correctly spelled single words, and these problems tend to carry over into their written composition (Connelly, Campbell, MacLean, & Barnes, 2006; Sterling, Farmer, Riddick, Morgan, & Matthews, 1998; Sumner, Connelly, & Barnett, 2013; Tops, Callens, van Cauwenberghe, Adriaens, & Brysbaert, 2013; Wengelin, 2002). Wengelin, for example, in a sample of Swedish adults producing multi-sentence expository texts found that 7.1% of words were spelled incorrectly by writers with dyslexia compared to near zero spelling errors in students who did not have a dyslexia diagnosis.

It is less clear to what extent spelling difficulties have knock-on effects for students' ability to generate higher level text features – sophisticated and syntactically-correct sentences, clear argumentative structure, idea-rich content, appropriate cohesion ties, and so forth. Mature writing is often assumed to be characterised by explicit, self-regulated thinking-and-reasoning (e.g., Bereiter, Burtis, & Scardamalia, 1988; Flower & Hayes, 1980; Hayes, 1996; but see Torrance, 2015). Several authors have argued that, by making excessive demands on processing capacity, spelling difficulty will constrain this higher-level processing (e.g., Berninger, 2000; Fayol, 1999; McCutchen, 1996). Writing, when proceeding fluently, is likely to rely on processing higher and lower level linguistic features in parallel (Olive, 2014) with, for example, ideas for the next sentence being retrieved, at least some of the time, while the preceding sentence is being executed. This parallelism breaks down if lower level processing becomes attention demanding, to the detriment of, for example, planning (or maintaining in working memory) what to write next.

This argument finds support in research that shows that spelling errors – specifically subject-verb agreement errors in French – increase under memory load not just in children but also in non-dyslexic adult writers (Fayol, Largy, & Lemaire, 1994; Totereau, Thevenin, & Fayol, 1997). Evidence that struggling with spelling has negative consequences for higher-level text feature is, however, mixed. Spelling ability uniquely predicts text quality in upper primary and lower secondary students with a diagnosis of dyslexia (Berninger, Nielsen, Abbott, Wijsman, & Raskind, 2008) although not in younger students without dyslexia (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997). Lower-primary students given specific spelling instruction show short-term gains in non-spelling aspects of performance on a sentence production task (Graham, Harris, & Chorzempa, 2002). However Berninger et al. (2002) found no benefits of teaching spelling over and above those afforded by teaching general composition skills.

Evidence for the effects of dyslexia in older students' writing is also patchy. Tops et al. (2013) compared summaries of a researcher-provided text written by Dutch university students with and without dyslexia – a task that requires both reading and writing skills. They found no differences in text length, lexical diversity[1], sentence length, or ratings of sentence

[1] Lexical diversity measured by type-token ratio: The number of unique words (word types) in the text divided by the total number of words (word tokens). High lexical diversity indicates texts that engage a broader vocabulary.

structure. However the texts of students with dyslexia had poorer structure and were less enjoyable to read. Hatcher, Snowling, and Griffiths (2002) also found poorer text structure, relative to controls, in English university students with dyslexia completing a similar task. Connelly et al. (2006) studied English university students writing essays on a familiar topic and with no reference texts to read. Ratings (of non-spelling-corrected texts) indicated poorer vocabulary and punctuation in students with dyslexia relative to age-matched controls, but no differences in ratings of higher-level text features (idea development, syntactic diversity, richness of content, or structure and coherence). Gajar (1989) found a similar pattern of results in a similar sample. By contrast, in a much younger dyslexic sample, Sumner et al. (2013) found that texts were shorter and of substantially poorer text quality, based on a composite measure including both structure and idea development, relative to age-matched controls.

There is some evidence, therefore, that struggling with decoding results in poorer quality text, particularly in younger writers. Smaller effects in older students, and particularly students who have made it as far as selective higher education (e.g., Connelly et al., 2006) may result from these students having adapted their writing processes or word choice to work around their spelling deficits. Sumner, Connelly, and Barnett (2014) found that in upper-primary students with dyslexia, but not age-matched controls, frequency of within-word pausing (periods when the pen was lifted for 30 *ms* or more) during a sentence-copying task was strongly correlated with spelling ability. This suggests that spelling difficulty, above a certain threshold, affects not just spelling accuracy but also the flow of production. It may be that this production disfluency – or, more precisely, the diversion of attention that caused this disfluency – rather than the tendency to make spelling errors *per se*, results in the production of poorer quality text. If a student manages to avoid the disruption caused by struggling with spelling a word when in mid flow – perhaps by delaying attending to possible errors until the sentence is complete – then we would not expect their difficulty with spelling to have negative consequences for other features of their text.

Understanding how decoding and spelling difficulties affect writing therefore requires detailed analysis of the extent to which students with and without decoding difficulties pause before or within words or make within-word edits. Wengelin (2007) counted frequency of pauses longer than two seconds, in a small sample of adults with dyslexia composing several different texts by keyboard. Compared with controls, pausing before words was twice as frequent, and within words over ten times as frequent (12% of words, compared to 1% for controls). Overall number of edits did not vary across groups, but editing to correct an error was much more common for students with dyslexia (27% compared to 2%). Texts produced by students with dyslexia showed lower lexical diversity and lexical density[2] than controls, with no difference between groups in a spoken-monologue condition. Frequency of within-word pausing and of editing both correlated strongly with lexical diversity and with lexical density in the dyslexic sample. This suggests a "word-level focus" in writers with dyslexia, with attention diverted from higher-level text features to producing correctly-spelled words. The word-level-focus hypothesis finds support in Sumner et al. (2013) who found no differences between 9 year olds with and without dyslexia in production speed when alphabet writing – a task that involves handwriting and recall but not spelling retrieval. However, on a composition task students with dyslexia wrote substantially shorter texts and, as we noted

[2] Lexical density is the proportion of content (open-class) words in the text and, broadly, indicates how much new information is provided by each sentence. Speech tends to be less lexically dense than text. In the present context, open class words comprise nouns, non-auxiliary verbs, adjectives, and adverbs.

above, made many more spelling errors. Slower production was associated not with slower handwritten production *per se* but with more hesitation (measured again as frequency of pen-lifts of > 30 *ms* duration).

These two studies may suggest that processing associated with word production, in students who struggle with decoding and / or spelling, demands active search of memory and other central processing mechanisms (e.g., Pashler, Johnston, & Ruthruff, 2001) which might otherwise be used for higher-level linguistic or semantic processes. Writers who do not struggle with spelling will typically produce words without this resulting in competition for the mechanisms required by these higher-level processes: Spelling retrieval occurs as part of a fluent cascade of low-level processing, leaving higher level functions to make use of central resources (e.g., Olive, 2014). In writers who do struggle with spelling, however, retrieval or assembly of the spelling of a word will require central resources, and therefore fluent, cascaded production breaks down and the writer hesitates. This account is consistent with increased within-word pausing in writers with dyslexia. It is also consistent with writers with dyslexia producing text with poorer structure and idea development, although this finding is open to alternative explanations.

Assuming, for the moment, that there is a word-level focus in the writing processes of students with dyslexia, it is worth asking why this occurs. There are (at least) two mutually compatible possibilities. The word-level focus may be primarily a production issue. Producing individual words is resource-demanding because the writer must, for a significant proportion of words, deliberately and explicitly remember how a word is spelled before they can output it on the page (or screen). This would be associated with pre-word and possibly mid-word hesitation, particularly for words that are then spelled incorrectly. Alternatively, word-level focus might primarily be associated with monitoring what appears on the page (a reading issue). Writers may read the word during or just after production and experience difficulty decoding what they have written: the word “doesn’t look right”. Analysis of the key-press data from writers with dyslexia reveals situations where words, including some correctly spelled words, are written but then immediately receive multiple edits as the writer struggles to produce a word that, to them, reads correctly (Wengelin, 2002), although we are not aware of evidence that this occurs frequently. This “monitoring hypothesis” predicts more hesitation mid-word and immediately post-word in weak decoders, relative to controls. It also predicts that if monitoring were to be prevented – if writers could not see the words that they are writing – then the phenomenon would disappear, resulting in less mid- and post-word hesitation, and possibly in benefits for higher-level features of the text.

The present study contributes to the surprisingly limited literature exploring the processes by which students with decoding difficulties (and students with dyslexia more generally) produce extended text. We compared the writing processes and written products of Norwegian upper-secondary students with poor decoding skills with controls matched on age and mathematics performance. Participants wrote by keyboard – their typical writing method. Norwegian is a shallow orthography (i.e. it shows a relatively regular phoneme to grapheme mapping). Our sample therefore contrasts with previous process-focussed studies, cited above, that have sampled either older and academically-selected students writing in a shallow orthography (Wengelin, 2007) or younger students writing in a language in which phoneme-grapheme mapping is particularly complex (Sumner et al., 2013).

Within this specific context we address three questions. First, do students who struggle with decoding produce poorer quality text? We answered this question by analysis of reader-based (holistic) ratings of participants’ texts, and also measures of spelling accuracy, word frequency, lexical diversity and density. This provided both an indication of whether text quality – as typically assessed in academic contexts – is poorer for weak decoders, but also of which specific text features are affected by decoding deficit. We predicted more spelling errors in weak decoders' texts. The theory that word-level focus reduces processing resources available for higher level processes specifically predicts that weak decoders produce text that shows some combination of weaker thematic development, poorer structure, narrower vocabulary, and text that is less lexically dense (more speech-like).

Second we tested the *word-level-focus hypothesis*: Do students who struggle with decoding devote disproportionate resources to processing at the word level? We examined the time-course of students' writing processes as revealed through analysis of the duration of intervals between key-presses – henceforth "key-press latencies". These analyses are based on the assumption that latency prior to output of, for example, a word or sentence, is a measure of the planning and retrieval or assembly processes associated with the word or sentence's production (see, for example, Damian & Stadthagen-Gonzalez, 2009; Maggio, Lété, Chenu, Jisa, & Fayol, 2012). We predicted a tendency towards longer latencies immediately before words, and perhaps also within words, although experimental evidence is mixed as to the extent that lexical and orthographic retrieval effects persist beyond production onset (Bertram, Tønnessen, Strömqvist, Hyönä, & Niemi, 2015; Bonin, Malardier, Méot, Fayol, & Meot, 2006; but see Damian & Stadthagen-Gonzalez, 2009; Delattre, Bonin, & Barry, 2006). We also examined effects specifically associated with generating and correcting spelling errors. Are errors associated with longer pre- and/or within-word latencies? Disfluency may also occur as a result of editing behaviour (stopping to make word-level changes). When spelling errors occur, do they tend to be detected and corrected?

Third we tested the *monitoring hypothesis*: Does word-level focus (if present) result from students experiencing decoding problems when reading the word they are currently producing or have just completed? We contrasted normal composition with a "masked text" condition in which students were prevented from reading what they were writing: In all analysis – of both written product and writing processes - comparison between the normal and masked-text conditions determined whether the extent to which differences between weak-decoders and controls can, in part, be attributed directly to weak decoders' difficulty in reading the text that they are writing or have just written.

Method

Design

Students identified as weak decoders and matched controls wrote short expository essays by keyboard using a simple text editor in each of two conditions. In the *normal* condition students could see what they were writing and had written. In the *masked-text* condition, each typed letter appeared on the screen as an X, thus preventing reading. Punctuation and spaces remained visible. Actual letters typed were recorded for subsequent analysis, and participants were aware that this was the case. Key presses (down strokes) – both what key was pressed and a precisely-recorded key-press time – were logged throughout. Dependent variables comprised text-based and reader-based assessment of written products and various process measures derived from analysis of key-press logs.

Participants

Our sample comprised 26 weak decoders and 26 controls, matched pairwise for age, mathematics performance and sex (26 female overall). Students were from upper secondary years (weak decoders, *M* = 16 years 11 months, *SD* = 9.6 months; controls, *M* =16 years 11 months, *SD* = 11.5 months) and were drawn from 12 rural and urban upper-secondary schools in western Norway.

Candidates for the weak decoders sample were initially identified by a combination of information from the Norwegian Pedagogical Psychology Service indicating a dyslexia diagnosis and by asking teachers to identify students within their classes who struggled with reading. These students were then tested on a standardized, Norwegian version of the word-split task (Jacobson, 2001; Miller-Guron, 1999) administered by the researchers. Word-split provides a fluency-focused measure of decoding ability which gives better discrimination than single-word reading accuracy measures in languages with transparent orthography (Wimmer, 1993). Each weak decoder was matched with a control student from the same class group, such that (a) minimum word-split score in the control condition was higher than that for the highest scoring weak decoder and (b) students were matched for age (born within the same academic year), mathematics performance (same or adjacent scores on a 6 point scale), and sex. Word-split means were below the 15th national-norm centile for weak decoders and above the 60th for controls (M = 36.9, SD = 12.3 and M = 59.6, SD = 6.4 respectively).

All weak decoders had, at some time in their school careers, received a formal dyslexia diagnosis based on performance on nationally standardized tests. Diagnosis requires listening comprehension within normal range and scores below the 30th centile on tests of at least three of the following: reading fluency, word identification, phonological awareness, word and non-word reading, and spelling. Diagnosis was also contingent on reading difficulties persisting despite conventional instruction and being resistant to intervention.

Mathematics performance was assessed by curriculum tasks distributed through the previous academic year and indexed against national standards (weak decoders, M = 3.00, SD = 1.01; control, M = 3.35, SD = 1.11; national mean = 3.5). This provided a matching criterion that captured academic performance that was sustained and self-regulated but that made relatively light demands on reading and writing ability. The present task – producing extended spontaneous text – is highly motivation- and prior-learning-dependent. Mathematics curriculum performance therefore provided a more informative matching criterion than would be provided by, for example, performance on a single non-verbal IQ test.

The Norwegian context

Norwegian spelling has largely regular phoneme-grapheme mapping and, with a few morphological exceptions, words can be accessed and spelled phonetically (Hagtvet, Helland, & Lyster, 2005, p. 21). Students diagnosed with dyslexia in Norway are typically assigned lessons with a specially trained teacher each week. These are typically reading-focussed but can also involve spelling practice and occasionally written composition. Students diagnosed with dyslexia are typically provided with a personal computer and assistive software to support spelling and text planning. From 8th grade all examinations and most other major assignments in Norwegian schools are usually written on computers. From the start of upper-secondary school all students are required to have their own laptop.

Writing tasks

Students wrote two expository texts, with order and topic counterbalanced across the normal / masked-text condition. Task statements were "*Bør lekser avskaffes? Skriv en sammenhengende tekst der du argumenterer for og mot dette.*" ("Should teachers cease to give students homework? Write an essay arguing the pros and cons of giving students homework.") and *"Bør Ungdomskortet være gratis? Skriv en sammenhengende tekst der du argumenterer for og mot dette."* ("Should public transportation for young people be free? Write an essay arguing the pros and cons of having free public transportations for young people"). Students wrote under "examination conditions" in a quiet room in their own school. They were tested in matched pairs. This provided some control over task duration without imposing a task-duration expectation across the whole sample (with negative consequences for the validity of the process data). Students were asked to approach the task as they would a normal in-class writing assignment. They were told that the person testing them, together with another researcher, would read their text. They were encouraged to do their best. Before writing in the masked text condition students were assured that writing in this condition is possible and what they wrote, although not visible, was stored.

Measures

Text quality. Quality measures were adapted from assessment criteria for the WIAT-II UK essay task (Wechsler, 2006). Texts were spelling-corrected before rating but punctuation and grammatical errors were preserved. They were scored for organization, theme development and vocabulary. *Organization* (maximum = 17) was assessed based on criteria including sentence and text structure, the presence of a topic sentence, whether the sentences in each paragraph followed a logical order, paragraph order and signposting. *Theme Development* (maximum = 8) was scored on the basis of the number and quality of supporting and counter arguments, maintenance of focus, and clear statement of position. *Vocabulary* (maximum = 7) was assessed by evaluating the appropriateness and diversity of the students' word choice.

All texts were scored independently on all three dimensions by two trained raters. Cohens weighted Kappa indicated good inter-rater reliability (Organization, .87; Theme Development, .85; Vocabulary, .82).

Text-based assessment of final texts. We recorded, for final texts, text length (number of word tokens), number of words incorrectly spelled, type-token ratio (as a measure of lexical diversity), ratio of open-class to closed-class words (as an indication of informational density), and mean word length and frequency (just for open-class words). Word frequency was calculated from the Norwegian Newspaper Corpus (NNC, 2013).

Process measures. Key-presses were logged using in-house software (anonymised, 20XX) implemented within the SR Research Experiment Builder environment. Our analyses are based on key-press latency. This refers to the interval between pressing the current key and the immediately preceding key, recorded in milliseconds and illustrated thus:

$$^{945}\mathrm{T}^{130}\mathrm{h}^{112}\mathrm{e}^{117}\_^{249}\mathrm{b}^{119}\mathrm{o}^{102}\mathrm{y}^{95}\_^{340}\mathrm{s}^{120}\mathrm{w}^{140}\mathrm{a}^{102}\mathrm{m}^{110}.$$

The example shows a pre-sentence latency of 945 *ms*, followed by within-word latencies of 130 *ms* for *h*, 112 *ms* for *e*. *Boy* has a pre-word latency of 249, and so forth. We assume that latency at a particular location is associated, in part, with cognitive activity associated with planning the linguistic unit that follows. Hence, for example, longer latencies before words than within words.

Results

We first report findings relating to the students' completed texts, and then explore the processes by which these were produced.

Written products

Table 1 summarizes findings from comparisons of texts written by weak decoders and control students in the normal and masked-text conditions. Inferential testing was by two way (group (weak decoders vs. control) by condition (masked text vs. normal text) mixed ANOVA, for analyses of quality and text-length effects, and by ANCOVA, with text length as a within-subjects covariate, for all other effects.

[insert Table 1 near here]

Ratings of the quality of students' spelling-corrected texts were significantly lower for weak decoders on all quality dimensions (Main effects of group: organization, $F(1,50) = 21$, $p < .001$, $\eta^2 = .30$; theme development $F(1,50) = 18$, $p < .001$, $\eta^2 = .27$; vocabulary $F(1,50) = 26$, $p < .001$, $\eta^2 = .36$) . In weak decoders only, masking the text resulted in poorer organization, although this effect was small ($F(1,50) = 4.8$, $p = .033$, $\eta^2 = .08$ for the group-by-condition interaction). We found no other statistically significant effects.

Weak decoders produced shorter texts than controls ($F(1,50) = 5.4$, $p = .024$, $\eta^2 = .10$ for the main effect of group). We found no effects involving condition.

Statistical analysis for the remaining product measures statistically controlled for text length, as detailed above. Number of spelling errors and word frequency were log transformed prior to analysis. As predicted, weak decoders made many more spelling errors than controls ($F(1,47) = 51$, $p < .001$, $\eta^2 = .47$ for the main effect of group). As might be expected, there were also reliably more spelling errors in the masked condition ($F(1,47) = 25$, $p < .001$, $\eta^2 = .32$ for the main effect of condition). The group-by-condition interaction was not significant. Weak decoders produced text with a lower ratio of open-class to closed class words, suggesting less informationally-rich text, although this effect was small ($F(1,47) = 8.3$, $p = .006$, $\eta^2 = .07$ for the main effect of group). Weak decoders also tended to use slightly shorter open-class (content) words ($F(1,47) = 11.6$, $p = .001$, $\eta^2 = .30$ for the main effect of group). We found no effects involving condition, and no significant effects for word frequency or type-token ratio.

Writing processes

In reporting process measures we first report total time on task. We then examine differences in key-press latency associated with location within the text (start of a sentence, start of a word, within a word, end of word). We then focus on just word-level effects, looking first at the probability of within-word editing, and finally at effects of spelling accuracy on key-press latency.

With the exception of analysis of total time on task, data comprised key-presses, nested within writing task, nested within student. Statistical inference involved testing linear mixed effects models, starting with a zero (intercept-only) model with random by-subject intercepts and random by-subject slopes for writing condition (masked vs. normal). We used χ^2 tests based on change in log-likelihood to determine whether differences in model fit were statistically significant. Statistical significance of parameters was established by evaluating against the z distribution.

Total time on task. Time on task is reported in Table 2. We found no statistically significant effect of group. We did, however, find an main effect of condition: Masking the text resulted in slightly less time-on-task in both groups ($F(1,50) = 7.1$, $p = .01$, $\eta^2 = .11$ for the main effect of condition in a group-by-condition ANOVA).

[insert Table 2 near here]

Key-press latencies. Key-press latencies were strongly negatively skewed. This skew exists in all response time data. However in the present context atypically long latencies cannot necessarily be dismissed as outliers: There is potential for a small number of long pauses to result from, or cause, substantial disruption to ongoing production. We therefore handle skew in two ways. In the first analysis that we report we trimmed to remove keystroke latencies longer than 8 s and then log-transformed prior to inferential testing[3]. In a second

[3] We checked the effects of removing latencies longer than 8 s by repeating the analysis on untrimmed data. There was no difference in pattern of effects, except in the sentence initial position which showed by-group differences in the untrimmed data. There were,

analysis, for word-level pauses only, we explored the frequency of pauses in several latency (pause duration) bins.

Table 3 gives keystroke latencies at six different locations within the text, by group and by condition (masked-text vs. normal). We tested four nested models, starting with an intercept-only model with random by-participant intercepts, and random by-participant slopes for both location and condition. Each successive model gave better fit (respectively, $\chi^2(7) > 100$, $p < .001$; $\chi^2(11) > 100$, $p < .001$; $\chi^2(5) = 12.1$, $p = .03$). We followed this analysis with three simultaneous linear contrasts conducted separately at each level of location (Table 3, lower panel).

[insert Table 3 near here]

Main findings were as follows: Within-word latencies were on average 27% (47 *ms*) longer for weak decoders. Masking the text slowed within-word latencies slightly for both groups. Pre-word latencies, which we assume are associated primarily with word-level planning, were 36% (115 *ms*) longer for weak decoders, consistent with the word-level-focus hypothesis. There was no effect of masking. Pre-sentence latencies – associated additionally with planning content and syntax – were substantially longer for weak decoders than for controls, and longer in both groups when the text was masked.

In a second analysis, we explored word-level effects by categorizing within-word and pre-word latencies into pause-length bins (.5 to 1 *s* , 1 to 2 *s*, 2 to 3 *s*, and > 3 *s*). We used logistic mixed-effects regression methods to predict the probability of bin membership on the basis of group and condition. We tested separate models at each pause-duration bin, with the binomial is-in-bin vs. is-not-in-bin as dependent variable. We fitted full factorial models with

---

proportionally, a much larger number of > 8 *s* latencies in the sentence initial position (M = 15%, SD = 19% across all students).

main effects for group (weak decoders vs. control) and condition (normal vs. masked), and their interaction, and random by-subject intercepts and condition slopes. Comparing fit of this model with that of the zero (intercept and random effects) model gave $\chi^2(3) = 9.2$, $p = .027$ for pre-word pauses in the range 2 to 3 *s*; $\chi^2(3) = 7.6$, $p = .055$ for 1 to 2 *s* mid-word pauses; and $\chi^2(3) > 13$, $p < .005$ in other cases except mid-word > 3 *s* pauses and pre-word 2 to 3 *s* pauses, neither of which gave significantly better fit.

Parameter estimates are reported in Table 4. Findings were as follows: Within word pausing in both .5 to 1 *s* and 1 to 2 *s* ranges was 2.6 times more probable in weak decoders than in controls. Pauses longer than 2 *s* were very rare in both groups. Pre-word pauses longer than 1 *s* were approximately twice as common in weak decoders, although this effect failed to reach significance for pauses in the 2 to 3 *s* range. Around 1 in 12 (non-sentence-initial) words written by weak decoders in the normal condition were preceded by pauses for longer than 2 *s*, compared to 1 in 26 for controls. Masking had no clear effect.

Relationship between word-level latencies and spelling accuracy. Here we focus just on within-word and word-initial latencies for words that were produced without internal editing. We determined the relationship between these latencies and the spelling accuracy of the resulting word. This addresses questions about whether writers who fail to spell a specific word correctly hesitate more prior to and/or during its production. We analyzed just open-class words to reduce confound with grammatical function. We also statistically controlled for word length and frequency. We modeled within-word and word-initial latencies separately, starting with a zero model, then adding word length and frequency as covariates, then main-effects of group, condition, and whether or not the word was correctly spelled, then two way interactions between these factors, and finally a full factorial model. For both within-word and before-word latencies the covariates-only model gave better fit that the zero model[4] ($\chi^2(2) > 100$, $p < .001$ for both within and before words). Adding main effects further improved fit ($\chi^2(3) = 103$, $p < .001$ within words; $\chi^2(3) = 36$, $p < .001$ before words), with significant effects for condition and group – repeating findings from the previous section – and a significant effect of whether or not the word was correctly spelled ($z = 8.3$, $p < .001$ within words; $z = 5.0$, $p < .001$ before words). We followed this with simultaneous pairwise contrasts, within groups, comparing correctly and incorrectly spelled words.

Estimated means and contrasts are given in Table 5. Findings were clear, consistent, and unexpected. In both groups words that were spelled incorrectly were associated with *shorter* preparation time before output was initiated and more rapid production when the word was being typed.

Word-level editing. We identified words as having been "internally edited" if at any point between the initial keypress and the keypress for the word-terminal character (space or punctuation) there were any delete or horizontal cursor-key presses[5] (henceforth *edit key-presses*). We tested logistic mixed-effects regression models, predicting the probability that a word was produced with internal editing on the basis of group and condition, first adding main effects to an intercept-only (zero) model, and then adding the interaction. These models showed significantly better fit (respectively, $\chi^2(2) = 34$, $p < .001$ and $\chi^2(1) = 4.7$, $p = .03$). As can be seen from Table 2, weak decoders in the seen-text condition were slightly less likely to edit words. This effect was not present when the text was masked.

[4] Latencies were longer for longer words and lower frequency words, both before and within words. $z > 6$, $p < .001$ in all cases.

[5] All cursor moves were by cursor key and all deletes by backspace. Vertical cursor-key presses, although potentially associated with editing, were omitted because this was also the mechanism by which text was vertically scrolled.

We then explored effects associated just with words that *were* edited during production. We found no evidence in the extent of editing across group and condition, measured in terms of edit keypresses. Spelling accuracy for edited words was slightly lower for weak-decoders (Table 2; $\chi^2(2) = 11.5$, $p = .003$ for the main effects model, as detailed above, with $z = 3.47$, $p = .001$ for the effect of group), but there were no main effects of condition or group-by-condition interaction. In response to a reviewer's suggestion, we also explored the possibility that, in the normal condition, editing took longer per keystroke for weak decoders, relative to what would be expected based on their normal (non-edited word) within-word production rate. We did not find evidence that this was the case.

Overall, we found no evidence that weak decoders were more likely to make changes to their text, based on the ratio of the number of edit keypresses to the number of character-key presses (Table 2). As might be expected, within-word editing was less common when the text was masked ($F(1,102) = 40$, $p < .001$ for a group x condition ANOVA). This effect was independent of group.

Discussion

Our study sampled students in mainstream upper-secondary education writing in a language with relatively shallow orthography. Within this specific context we aimed to answer three questions: 1. Do students who struggle with decoding produce poorer quality texts? 2. Is there evidence of a word-level focus in weak-decoder's writing processes? 3. If so, is this primarily a production deficit or does it result from problems reading what they are or have just written (the "monitoring hypothesis")? We will address each question in turn.

Weak decoders in the normal writing condition made, on average, more than four times as many spelling errors as controls. This is consistent with previous findings (Connelly et al., 2006; Sterling et al., 1998; Tops et al., 2013; Wengelin, 2002). Spelling aside, weak decoders produced text that was rated as having poorer use of vocabulary, poorer theme development, and being less well structured. Evidence for quality differences in this study was stronger than in the only previous study that we are aware of that has tested a broadly comparable sample and used a similar writing task (Connelly et al., 2006). Our findings are consistent with the hypothesis that spelling difficulties hamper students' ability to produce higher-level text features, although, as we suggest below, correlation does not necessarily imply causality. Third-factor explanations are equally plausible.

Findings from analysis of key-press latencies provided partial evidence for the word-level-focus hypothesis. Latencies were longer for weak decoders, relative to controls, at all text locations. Word-initial latencies, which are likely to be associated with retrieving spelling, showed clear effects both in mean duration and, perhaps more importantly, in the proportion of longer pauses. This suggests that weak decoders found word-level planning generally more effortful and experienced a larger number of interruptions to fluent production.

Word-level hesitation appeared, however, to be entirely associated with production rather than reading. According to the "monitoring hypothesis" disfluency results in part from students with dyslexia struggling to read the word that they are writing or have just written and/or identifying this word as misspelled. If weak decoders in our sample had spent time looking back at words they had just written and puzzling over whether or not they were correctly spelled, then we would expect differences between groups to disappear when they were prevented from seeing what they had written. This was not what we found. In fact in both groups word-end latencies were typically similar to or slightly shorter than within-word latencies. This suggests that the move from the word-terminal character to the space bar was not associated with any additional processing beyond that which occurred within-word. It is also possible that inflated word-initial latencies in weak decoders occurred because students looked back at the just-produced word (or earlier words). However if this were the case then the difference between groups should disappear in the masked-text condition. Again, this is not what we found. More generally, in the context of the relatively short writing task used in this study, we found that masking the text, and therefore preventing reading what was written, did not have any negative consequences for text quality in either group. This result is perhaps counterintuitive, but it is consistent with previous findings (Gould & Boise, 1978; Hull & Smith, 1983; but see Olive & Piolat, 2002).

Compared to controls, weak decoders were somewhat slower in typing words, after the initial character. They were, however, no more likely than controls to engage in within-word editing and there was no evidence that when they did edit they did so more extensively. The writing behavior of weak decoders in our sample therefore was not consistent with anecdotal accounts of very extensive and repeated within-word delay and editing by writers with dyslexia. Failure to retrieve an accurate spelling was, in fact, associated with *shorter* latencies, in both groups, both before and within words. Thus, although weak decoders found retrieval of correct spelling more effortful relative to controls, there was no evidence that failure to retrieve correct spelling typically introduced further disfluency. On the contrary, production tended to be faster when spelling was incorrect.

Our results suggest, therefore, that weak decoders (and, by extension, students with a dyslexia diagnosis) tend to make more spelling errors and produce text that was judged as lexically, structurally, and thematically weaker than that of peers with average decoding ability, even after spelling was corrected. Our results also suggest that, for spelling rather than decoding reasons, weak decoders take longer, and are more likely to hesitate, when preparing words. What is less clear is whether these effects are causally related. Is it the case, as we hypothesized in our introduction, that poorer text in weak decoders was caused by diversion of attention to the word level reducing resources available for the higher-level processes necessary to create well-formed text? Some evidence in support of this argument perhaps comes from the fact that we found substantially longer sentence-initial latencies in weak decoders. In experiments exploring written sentence production, Nottbusch (2010) found sentence-initial latencies of around 1300 *ms* (timed between subjects first seeing a stimulus array and the first keystroke of their response). This value is only slightly higher than the sentence-initial latencies found for weak-decoders in this study, but much higher than for controls. It may be that control students were able to start planning the content and syntax of their next sentence in parallel with typing. The higher demands of word-level processing for weak-decoders mean that in these students sentence planning was partly or entirely delayed until the inter-sentence pause.

However, even if word-level focus reduces parallelism in production, it is not clear that this will necessarily have consequences for higher-level text features. It is equally plausible that relatively weak structure and content are due to effects of reading deficit on prior learning (see Stanovich, 1994, for an analogous argument concerning the relationship between reading and verbal ability). The primary deficit associated with poor decoding skills (and therefore a probable dyslexia diagnosis) is difficulty with reading. This results in both a poorer comprehension (Ransby & Swanson, 2003) and a general tendency to read less (Mol & Bus, 2011). During their school careers students with dyslexia will, therefore, be less well placed to learn, through reading, the linguistic strategies and conventions that will support their own production of coherent text. This, rather than struggling with spelling when writing, may be a more parsimonious explanation for weak decoders' relatively poor text.

In conclusion, therefore, our findings suggest that struggling with decoding has negative consequences for students' written products and has measurable effects on writing processes. We did not, however, find evidence of the kinds of extensive within-word editing or frequent very long pauses that would be strong evidence that poorer quality text results directly from a word-level focus. Whether or not stronger effects would emerge with, for example, a younger sample writing in a less transparent orthography is a question that could usefully be explored in future research.

Acknowledgements

References

Bereiter, C., Burtis, P., & Scardamalia, M. (1988). Cognitive operations in constructing main points in written composition. *Journal of Memory and Language*, *27*(3), 261–278. doi:10.1016/0749-596X(88)90054-X

Berninger, V., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., … Graham, S. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology*. doi:10.1037/0022-0663.94.2.291

Berninger, V. W. (2000). Dyslexia the invisible, treatable disorder: The story of Einstein's ninja turtles. *Learning Disability Quarterly*, *23*(3), 175–195. doi:10.2307/1511163

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: under-recognized and under-treated. *Journal of School Psychology*, *46*(1), 1–21. doi:10.1016/j.jsp.2006.11.008

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), 1–10. doi:10.3389/fnhum.2015.00207

Bonin, P., Malardier, N., Méot, A., Fayol, M., & Meot, A. (2006). The scope of advance planning in written picture naming. *Language and Cognitive Processes*, *21*(1-3), 205–237. doi:10.1080/01690960400002141

Connelly, V., Campbell, S., MacLean, M., & Barnes, J. (2006). Contribution of Lower Order Skills to the Written Composition of College Students With and Without Dyslexia. *Developmental Neuropsychology*, *29*(1), 175–196. doi:10.1207/s15326942dn2901_9

Damian, M., & Stadthagen-Gonzalez, H. (2009). Advance planning of form properties in the written production of single and multiple words. *Language and Cognitive Processes*, *24*(4), 555–579. Retrieved from <Go to ISI>://000265359600004

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

Fayol, M. (1999). From on-line management problems to strategies in written composition. In M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing: processing capacity and working memory effects in text production* (pp. 15–23). Amsterdam: Amsterdam University Press.

Fayol, M., Largy, P., & Lemaire, P. (1994). A special case of agreement error in French: The confusion of noun/verb inflections. In G. Rijlaarsdam, H. van den Burgh, & M. Couzijn (Eds.), *Theories, models and methodology in writing research* (pp. 169–212). Amsterdam: Amsterdam University Press.

Flower, L., & Hayes, J. (1980). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp.

31–50). Hillsdale, N.J.: Erlbaum.

Gajar, A. H. (1989). A computer analysis of written language variables and a comparison of compositions written by university students with and without learning disabilities. *Journal of Learning Disabilities*, *22*(2), 125–130. doi:10.1177/002221948902200208

Gould, J. D., & Boise, S. J. (1978). How authors think about their writing, dictating and speaking. *Human Factors*, *20*, 494–505.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997, March). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*. doi:10.1037/0022-0663.89.1.170

Graham, S., Harris, K. R., & Chorzempa, B. F. (2002, December). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology*. doi:10.1037/0022-0663.94.4.669

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, N.J.: Erlbaum.

Hatcher, J., Snowling, M. J., & Griffiths, Y. M. (2002). Cognitive assessment of dyslexic students in higher education. *The British Journal of Educational Psychology*, *72*(Pt 1), 119–133. doi:10.1348/000709902158801

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–26). Mahaw, N.J.: Erlbaum.

Hull, G. a, & Smith, W. L. (1983). Interrupting visual feedback in writing. *Perceptual and Motor Skills*, *57*(3 Pt 1), 963–978. doi:10.2466/pms.1983.57.3.963

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

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*. doi:10.1007/s11881-003-0001-9

Maggio, S., Lété, B., Chenu, F., Jisa, H., & Fayol, M. (2012). Tracking the mind during writing: Immediacy, delayed, and anticipatory effects on pauses and writing rate. *Reading and Writing*, *25*, 2131–2151. doi:10.1007/s11145-011-9348-1

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

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

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267–296. doi:10.1037/a0021890

Nottbusch, G. (2009). Grammatical planning, execution, and control in written sentence production. *Reading and Writing*, *23*(7), 777–801. doi:10.1007/s11145-009-9188-4

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

Olive, T., & Piolat, A. (2002). Suppressing visual feedback in written composition: Effects on processing demands and coordination of the writing processes. *International Journal of Psychology*, *37*(4), 209–218. doi:10.1080/00207590244000089

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology*, *52*, 629–651. doi:10.1146/annurev.psych.52.1.629

Ransby, M. J., & Swanson, H. L. (2003). Reading comprehension skills of young adults with childhood diagnoses of dyslexia. *Journal of Learning Disabilities*, *36*(6), 538–555.

doi:10.1177/00222194030360060501

Stanovich, K. S. (1994). Annotation: Does Dyslexia Exist. *Journal of Child Psychology and Psychiatry*, *35*(4), 579–595. doi:10.1111/j.1469-7610.1994.tb01208.x

Sterling, C., Farmer, M., Riddick, B., Morgan, S., & Matthews, C. (1998). Adult Dyslexic Writing. *Dyslexia*, *4*, 1–15. doi:D-DYS87>3.0.CO;2-F

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

Sumner, E., Connelly, V., & Barnett, A. L. (2014). The influence of spelling ability on handwriting production: children with and without dyslexia. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(5), 1441–7. doi:10.1037/a0035785

Tops, W., Callens, C., van Cauwenberghe, E., Adriaens, J., & Brysbaert, M. (2013). Beyond spelling: The writing skills of students with dyslexia in higher education. *Reading and Writing*, *26*(5), 705–720. doi:10.1007/s11145-012-9387-2

Torrance, M. (2015). Understanding Planning in Text Production. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (2nd Editio.). New York: Guildford Press.

Totereau, C., Thevenin, M., & Fayol, M. (1997). The development of the understanding of number morphology in written French. In C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to Spell* (pp. 97–114). Hillsdale, NJ: Erlbaum.

Wechsler, D. (2006). *Wechsler Individual Achievement Test - Second UK Edition (WIAT-II UK)*. Oxford: Pearson Assessment.

Wengelin, Å. (2002). *Text production in adults with reading and writing difficulties (Gothenburg Monographs of Linguistics, 20)*. Gothenburg: Gothenburg University.

Wengelin, Å. (2007). The word level focus in text production by adults with reading and writing difficulties. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and Cognition: Research and Applications* (pp. 68–82). Amsterdam: Elsevier.

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

Figures and Tables

Table 1. Text assessment and global process measures. Mean values with standard deviation in parenthesis.

	Weak Decoders		Controls	
	Normal	Masked	Normal	Masked
Quality ratings				
Organisation**	6.92 (3.51)	5.85 (3.16)	10.27 (3.44)	10.30 (2.96)
Theme development**	4.96 (1.25)	5.04 (1.18)	6.12 (.82)	6.07 (1.04)
Vocabulary**	1.42 (1.33)	1.42 (1.14)	3.04 (1.40)	2.93 (1.11)
Text-based measures				
Text length (words)*	257 (88)	246 (96)	307 (115)	341 (175)
Incorrectly spelled words**	18.3 (14.9)	20.1 (9.8)	4.0 (4.7)	8.1 (6.5)
Open-class to closed-class ratio**	1.24 (.20)	1.17 (.26)	1.34 (.22)	1.37 (.20)
Type-token ratio	.61 (.07)	.63 (.07)	.62 (.09)	.60 (.07)
Mean word frequency	1098 (276)	1102 (343)	937 (150)	929 (212)
Mean word length**	5.4 (.4)	5.3 (.5)	5.8 (.3)	5.7 (.2)

Note: Type-token ratio, frequency and length are just for open class words. Frequency per million. Group (Weak Decoders vs. Controls) main effect, * $p < .05$, ** $p < .01$

Table 2. Writing process measures by condition and group. By-subject means, with *SD* in parenthesis.

	Weak Decoders		Controls	
	Normal	Masked	Normal	Masked
Total time on task (minutes, seconds)	19,56 (7,18)	16,30 (7,15)	20,11 (9,17)	18,52 (10,12)
Proportion of words that were produced with internal editing[a]	.10 (.04)	.07 (.03)	.14 (.04)	.08 (.04)
Proportion of internally-edited words that were then correctly spelled**	.87 (.10)	.85 (.14)	.92 (.05)	.94 (.08)
Editing ratio	.51 (.39)	.17 (.13)	.46 (.20)	.22 (.11)

Note. Editing ratio = total number of delete-key and horizontal cursor-key presses divided by number of character-key presses. Group (Weak Decoders vs. Controls) main effects, * $p < .05$, ** $p < .005$; [a]Group x condition interaction, $p = .03$.

Table 3. Keystroke latencies (*ms*) at different locations within the text. Parameter estimates, with 95% CI in parenthesis. Lower panel reports contrast tests (*z*, *p*)

	Within word *b^o^y*	Word-end *The^ boy^*	Pre-word *The ^boy*	Pre-sentence *swam. ^Then*
Weak Decoders: Normal	214 (197, 232)	184 (169, 200)	441 (406, 479)	1042 (917, 1183)
Weak Decoders: Masked	224 (207, 243)	195 (180, 212)	425 (392, 461)	1232 (1088, 1394)
Control: Normal	167 (154, 181)	155 (143, 168)	321 (296, 348)	716 (641, 800)
Control: Masked	180 (166, 194)	164 (152, 178)	315 (291, 341)	913 (824, 1011)
Weak Decoders vs. Control	4.1, < .001	3.0, .003	5.4, < .001	4.7, < .001
Weak Decoders: Normal vs. Masked	2.7, .006	3.2, .003	1.7, *ns*	2.4, .019
Control: Normal vs. Masked	4.6, < .001	3.2, .003	.20, *ns*	4.5, < .001

Note: Italicised text illustrates boundary (indicated by ^). Contrasts control for familywise error rate within text-boundary type. *ns* for *p* > .05

Table 4. Proportion of pauses, of various durations, by group (weak decoders vs. control) and condition (normal vs. masked text). Parameter estimates, with 95% CI in parenthesis

	.5 to 1 *s*	1 to 2 *s*	2 to 3 *s*	> 3 *s*
Within word pauses				
Weak Decoders: Normal	.059 (.046, .074)	.010 (.008, .014)	.001 (.001, .002)	.000 (.000, .001)
Weak Decoders: Masked	.054 (.043, .066)	.011 (.008, .016)	.001 (.000, .001)	.001 (.000, .001)
Control: Normal	.023 (.018, .030)	.006 (.004, .008)	.000 (.000, .001)	.000 (.000, .001)
Control: Masked	.026 (.021, .032)	.005 (.003, .006)	.001 (.000, .001)	.000 (.000, .001)
Effects[a]				
group	4.7, < .001	3.8, < .001	.05, *ns*	1.7, *ns*
condition	1.7, *ns*	.81, *ns*	1.5, *ns*	1.8, *ns*
interaction	2.4, .02	1.7, *ns*	2.3, *ns*	1.4. *ns*
Pre-word pauses				
Weak Decoders: Normal	.180 (.143, .224)	.085 (.068, .106)	.025 (.019, .033)	.032 (.024, .043)
Weak Decoders: Masked	.203 (.164, .248)	.085 (.065, .111)	.025 (.019, .033)	.039 (.029, .052)
Control: Normal	.128 (.101, .161)	.054 (.043, .067)	.017 (.013, .023)	.018 (.013, .023)
Control: Masked	.155 (.125, .191)	.052 (.039, .068)	.018 (.013, .024)	.017 (.012, .023)
Effects[a]				
group	1.8, *ns*	2.5, .011	1.7, *ns*	4.0, < .001
condition	2.3, .019	.01, *ns*	.06, *ns*	1.7, *ns*
interaction	.85, *ns*	.25, *ns*	.12, *ns*	1.6, *ns*

Note. [a]Effects (*z, p*) from logistic mixed effects regression models tested separately for each pause duration at each location, modelling the probability that pause was of that duration. *ns* for *p* > .05

Table 5. Keystroke latencies (*ms*) by whether or not the word was correctly spelled. Parameter estimates, with 95% CI in parenthesis.

Location	Spelling	Weak Decoders		Control	
		Normal	Masked	Normal	Masked
Within word	Correct	225 (208, 243)	235 (217, 254)	170 (157, 183)	180 (167, 195)
	Incorrect	206 (189, 224)	211 (193, 230)	148 (135, 162)	171 (157, 187)
	Correct vs. Incorrect[a]	4.8, < .001	6.0, < .001	5.3, < .001	2.3, .019
Pre-word	Correct	502 (430, 585)	459 (399, 528)	358 (308, 416)	352 (307, 403)
	Incorrect	392 (326, 472)	406 (342, 482)	282 (227, 350)	302 (246, 370)
	Correct vs. Incorrect[a]	4.2, < .001	2.2, .030	2.9, .008	1.9, .056

Note: [a]*z* and *p* from linear contrast, controlling for familywise error rate within group.