



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

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

SCHOLARONE™
Manuscripts

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

1
2
3
4 likely to rely on processing higher and lower level linguistic features in parallel (Olive, 2014)
5
6 with, for example, ideas for the next sentence being retrieved, at least some of the time, while
7
8 the preceding sentence is being executed. This parallelism breaks down if lower level
9
10 processing becomes attention demanding, to the detriment of, for example, planning (or
11
12 maintaining in working memory) what to write next.
13

14
15 This argument finds support in research that shows that spelling errors – specifically
16
17 subject-verb agreement errors in French – increase under memory load not just in children
18
19 but also in non-dyslexic adult writers (Fayol, Largy, & Lemaire, 1994; Totereau, Thevenin, &
20
21 Fayol, 1997). Evidence that struggling with spelling has negative consequences for higher-
22
23 level text feature is, however, mixed. Spelling ability uniquely predicts text quality in upper
24
25 primary and lower secondary students with a diagnosis of dyslexia (Berninger, Nielsen,
26
27 Abbott, Wijsman, & Raskind, 2008) although not in younger students without dyslexia
28
29 (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997). Lower-primary students given
30
31 specific spelling instruction show short-term gains in non-spelling aspects of performance on
32
33 a sentence production task (Graham, Harris, & Chorzempa, 2002). However Berninger et al.
34
35 (2002) found no benefits of teaching spelling over and above those afforded by teaching
36
37 general composition skills.
38
39
40

41
42 Evidence for the effects of dyslexia in older students' writing is also patchy. Tops et al.
43
44 (2013) compared summaries of a researcher-provided text written by Dutch university
45
46 students with and without dyslexia – a task that requires both reading and writing skills. They
47
48 found no differences in text length, lexical diversity¹, sentence length, or ratings of sentence
49

50
51 ¹ Lexical diversity measured by type-token ratio: The number of unique words (word types)
52
53 in the text divided by the total number of words (word tokens). High lexical diversity
54
55 indicates texts that engage a broader vocabulary.
56

1
2
3
4 structure. However the texts of students with dyslexia had poorer structure and were less
5 enjoyable to read. Hatcher, Snowling, and Griffiths (2002) also found poorer text structure,
6 relative to controls, in English university students with dyslexia completing a similar task.
7
8
9
10
11 Connelly et al. (2006) studied English university students writing essays on a familiar topic
12 and with no reference texts to read. Ratings (of non-spelling-corrected texts) indicated poorer
13 vocabulary and punctuation in students with dyslexia relative to age-matched controls, but no
14 differences in ratings of higher-level text features (idea development, syntactic diversity,
15 richness of content, or structure and coherence). Gajar (1989) found a similar pattern of
16 results in a similar sample. By contrast, in a much younger dyslexic sample, Sumner et al.
17 (2013) found that texts were shorter and of substantially poorer text quality, based on a
18 composite measure including both structure and idea development, relative to age-matched
19 controls.
20
21
22
23
24
25
26
27
28
29

30
31 There is some evidence, therefore, that struggling with decoding results in poorer quality
32 text, particularly in younger writers. Smaller effects in older students, and particularly
33 students who have made it as far as selective higher education (e.g., Connelly et al., 2006)
34 may result from these students having adapted their writing processes or word choice to work
35 around their spelling deficits. Sumner, Connelly, and Barnett (2014) found that in upper-
36 primary students with dyslexia, but not age-matched controls, frequency of within-word
37 pausing (periods when the pen was lifted for 30 *ms* or more) during a sentence-copying task
38 was strongly correlated with spelling ability. This suggests that spelling difficulty, above a
39 certain threshold, affects not just spelling accuracy but also the flow of production. It may be
40 that this production disfluency – or, more precisely, the diversion of attention that caused this
41 disfluency – rather than the tendency to make spelling errors *per se*, results in the production
42 of poorer quality text. If a student manages to avoid the disruption caused by struggling with
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 spelling a word when in mid flow – perhaps by delaying attending to possible errors until the
5 sentence is complete – then we would not expect their difficulty with spelling to have
6 negative consequences for other features of their text.
7
8
9

10 Understanding how decoding and spelling difficulties affect writing therefore requires
11 detailed analysis of the extent to which students with and without decoding difficulties pause
12 before or within words or make within-word edits. Wengelin (2007) counted frequency of
13 pauses longer than two seconds, in a small sample of adults with dyslexia composing several
14 different texts by keyboard. Compared with controls, pausing before words was twice as
15 frequent, and within words over ten times as frequent (12% of words, compared to 1% for
16 controls). Overall number of edits did not vary across groups, but editing to correct an error
17 was much more common for students with dyslexia (27% compared to 2%). Texts produced
18 by students with dyslexia showed lower lexical diversity and lexical density² than controls,
19 with no difference between groups in a spoken-monologue condition. Frequency of within-
20 word pausing and of editing both correlated strongly with lexical diversity and with lexical
21 density in the dyslexic sample. This suggests a “word-level focus” in writers with dyslexia,
22 with attention diverted from higher-level text features to producing correctly-spelled words.
23 The word-level-focus hypothesis finds support in Sumner et al. (2013) who found no
24 differences between 9 year olds with and without dyslexia in production speed when alphabet
25 writing – a task that involves handwriting and recall but not spelling retrieval. However, on a
26 composition task students with dyslexia wrote substantially shorter texts and, as we noted
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 ² Lexical density is the proportion of content (open-class) words in the text and, broadly,
49 indicates how much new information is provided by each sentence. Speech tends to be less
50 lexically dense than text. In the present context, open class words comprise nouns, non-
51 auxiliary verbs, adjectives, and adverbs.
52
53
54
55
56
57
58
59
60

1
2
3
4 above, made many more spelling errors. Slower production was associated not with slower
5
6 handwritten production *per se* but with more hesitation (measured again as frequency of pen-
7
8 lifts of > 30 *ms* duration).
9

10
11 These two studies may suggest that processing associated with word production, in
12
13 students who struggle with decoding and / or spelling, demands active search of memory and
14
15 other central processing mechanisms (e.g., Pashler, Johnston, & Ruthruff, 2001) which might
16
17 otherwise be used for higher-level linguistic or semantic processes. Writers who do not
18
19 struggle with spelling will typically produce words without this resulting in competition for
20
21 the mechanisms required by these higher-level processes: Spelling retrieval occurs as part of
22
23 a fluent cascade of low-level processing, leaving higher level functions to make use of central
24
25 resources (e.g., Olive, 2014). In writers who do struggle with spelling, however, retrieval or
26
27 assembly of the spelling of a word will require central resources, and therefore fluent,
28
29 cascaded production breaks down and the writer hesitates. This account is consistent with
30
31 increased within-word pausing in writers with dyslexia. It is also consistent with writers with
32
33 dyslexia producing text with poorer structure and idea development, although this finding is
34
35 open to alternative explanations.
36
37
38

39
40 Assuming, for the moment, that there is a word-level focus in the writing processes of
41
42 students with dyslexia, it is worth asking why this occurs. There are (at least) two mutually
43
44 compatible possibilities. The word-level focus may be primarily a production issue.
45
46 Producing individual words is resource-demanding because the writer must, for a significant
47
48 proportion of words, deliberately and explicitly remember how a word is spelled before they
49
50 can output it on the page (or screen). This would be associated with pre-word and possibly
51
52 mid-word hesitation, particularly for words that are then spelled incorrectly. Alternatively,
53
54 word-level focus might primarily be associated with monitoring what appears on the page (a
55
56
57
58
59
60

1
2
3
4 reading issue). Writers may read the word during or just after production and experience
5
6 difficulty decoding what they have written: the word “doesn’t look right”. Analysis of the
7
8 key-press data from writers with dyslexia reveals situations where words, including some
9
10 correctly spelled words, are written but then immediately receive multiple edits as the writer
11
12 struggles to produce a word that, to them, reads correctly (Wengelin, 2002), although we are
13
14 not aware of evidence that this occurs frequently. This “monitoring hypothesis” predicts more
15
16 hesitation mid-word and immediately post-word in weak decoders, relative to controls. It also
17
18 predicts that if monitoring were to be prevented – if writers could not see the words that they
19
20 are writing – then the phenomenon would disappear, resulting in less mid- and post-word
21
22 hesitation, and possibly in benefits for higher-level features of the text.
23
24

25
26 The present study contributes to the surprisingly limited literature exploring the
27
28 processes by which students with decoding difficulties (and students with dyslexia more
29
30 generally) produce extended text. We compared the writing processes and written products of
31
32 Norwegian upper-secondary students with poor decoding skills with controls matched on age
33
34 and mathematics performance. Participants wrote by keyboard – their typical writing method.
35
36 Norwegian is a shallow orthography (i.e. it shows a relatively regular phoneme to grapheme
37
38 mapping). Our sample therefore contrasts with previous process-focussed studies, cited
39
40 above, that have sampled either older and academically-selected students writing in a shallow
41
42 orthography (Wengelin, 2007) or younger students writing in a language in which phoneme-
43
44 grapheme mapping is particularly complex (Sumner et al., 2013).
45
46
47

48
49 Within this specific context we address three questions. First, do students who struggle
50
51 with decoding produce poorer quality text? We answered this question by analysis of reader-
52
53 based (holistic) ratings of participants’ texts, and also measures of spelling accuracy, word
54
55 frequency, lexical diversity and density. This provided both an indication of whether text
56
57
58
59
60

1
2
3
4 quality – as typically assessed in academic contexts – is poorer for weak decoders, but also of
5
6 which specific text features are affected by decoding deficit. We predicted more spelling
7
8 errors in weak decoders’ texts. The theory that word-level focus reduces processing resources
9
10 available for higher level processes specifically predicts that weak decoders produce text that
11
12 shows some combination of weaker thematic development, poorer structure, narrower
13
14 vocabulary, and text that is less lexically dense (more speech-like).

15
16
17 Second we tested the *word-level-focus hypothesis*: Do students who struggle with
18
19 decoding devote disproportionate resources to processing at the word level? We examined the
20
21 time-course of students’ writing processes as revealed through analysis of the duration of
22
23 intervals between key-presses – henceforth “key-press latencies”. These analyses are based
24
25 on the assumption that latency prior to output of, for example, a word or sentence, is a
26
27 measure of the planning and retrieval or assembly processes associated with the word or
28
29 sentence’s production (see, for example, Damian & Stadthagen-Gonzalez, 2009; Maggio,
30
31 Lété, Chenu, Jisa, & Fayol, 2012). We predicted a tendency towards longer latencies
32
33 immediately before words, and perhaps also within words, although experimental evidence is
34
35 mixed as to the extent that lexical and orthographic retrieval effects persist beyond
36
37 production onset (Bertram, Tønnessen, Strömqvist, Hyönä, & Niemi, 2015; Bonin, Malardier,
38
39 Méot, Fayol, & Meot, 2006; but see Damian & Stadthagen-Gonzalez, 2009; Delattre, Bonin,
40
41 & Barry, 2006). We also examined effects specifically associated with generating and
42
43 correcting spelling errors. Are errors associated with longer pre- and/or within-word
44
45 latencies? Disfluency may also occur as a result of editing behaviour (stopping to make word-
46
47 level changes). When spelling errors occur, do they tend to be detected and corrected?
48
49
50
51
52

53 Third we tested the *monitoring hypothesis*: Does word-level focus (if present) result
54
55 from students experiencing decoding problems when reading the word they are currently
56
57
58
59
60

1
2
3
4 producing or have just completed? We contrasted normal composition with a “masked text”
5
6 condition in which students were prevented from reading what they were writing: In all
7
8 analysis – of both written product and writing processes - comparison between the normal
9
10 and masked-text conditions determined whether the extent to which differences between
11
12 weak-decoders and controls can, in part, be attributed directly to weak decoders’ difficulty in
13
14 reading the text that they are writing or have just written.
15
16

17 18 **Method**

19 20 21 **Design**

22
23 Students identified as weak decoders and matched controls wrote short expository essays by
24
25 keyboard using a simple text editor in each of two conditions. In the *normal* condition
26
27 students could see what they were writing and had written. In the *masked-text* condition, each
28
29 typed letter appeared on the screen as an X, thus preventing reading. Punctuation and spaces
30
31 remained visible. Actual letters typed were recorded for subsequent analysis, and participants
32
33 were aware that this was the case. Key presses (down strokes) – both what key was pressed
34
35 and a precisely-recorded key-press time – were logged throughout. Dependent variables
36
37 comprised text-based and reader-based assessment of written products and various process
38
39 measures derived from analysis of key-press logs.
40
41
42
43

44 45 **Participants**

46
47 Our sample comprised 26 weak decoders and 26 controls, matched pairwise for age,
48
49 mathematics performance and sex (26 female overall). Students were from upper secondary
50
51 years (weak decoders, $M = 16$ years 11 months, $SD = 9.6$ months; controls, $M = 16$ years 11
52
53 months, $SD = 11.5$ months) and were drawn from 12 rural and urban upper-secondary schools
54
55 in western Norway.
56

1
2
3
4 Candidates for the weak decoders sample were initially identified by a combination of
5 information from the Norwegian Pedagogical Psychology Service indicating a dyslexia
6 diagnosis and by asking teachers to identify students within their classes who struggled with
7 reading. These students were then tested on a standardized, Norwegian version of the word-
8 split task (Jacobson, 2001; Miller-Guron, 1999) administered by the researchers. Word-split
9 provides a fluency-focused measure of decoding ability which gives better discrimination
10 than single-word reading accuracy measures in languages with transparent orthography
11 (Wimmer, 1993). Each weak decoder was matched with a control student from the same class
12 group, such that (a) minimum word-split score in the control condition was higher than that
13 for the highest scoring weak decoder and (b) students were matched for age (born within the
14 same academic year), mathematics performance (same or adjacent scores on a 6 point scale),
15 and sex. Word-split means were below the 15th national-norm centile for weak decoders and
16 above the 60th for controls ($M = 36.9$, $SD = 12.3$ and $M = 59.6$, $SD = 6.4$ respectively).
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33 All weak decoders had, at some time in their school careers, received a formal dyslexia
34 diagnosis based on performance on nationally standardized tests. Diagnosis requires listening
35 comprehension within normal range and scores below the 30th centile on tests of at least
36 three of the following: reading fluency, word identification, phonological awareness, word
37 and non-word reading, and spelling. Diagnosis was also contingent on reading difficulties
38 persisting despite conventional instruction and being resistant to intervention.
39
40
41
42
43
44
45

46 Mathematics performance was assessed by curriculum tasks distributed through the
47 previous academic year and indexed against national standards (weak decoders, $M = 3.00$, SD
48 $= 1.01$; control, $M = 3.35$, $SD = 1.11$; national mean $= 3.5$). This provided a matching
49 criterion that captured academic performance that was sustained and self-regulated but that
50 made relatively light demands on reading and writing ability. The present task – producing
51
52
53
54
55
56
57
58
59
60

1
2
3
4 extended spontaneous text – is highly motivation- and prior-learning-dependent. Mathematics
5
6 curriculum performance therefore provided a more informative matching criterion than would
7
8 be provided by, for example, performance on a single non-verbal IQ test.
9

10 11 **The Norwegian context**

12
13 Norwegian spelling has largely regular phoneme-grapheme mapping and, with a few
14
15 morphological exceptions, words can be accessed and spelled phonetically (Hagtvet, Helland,
16
17 & Lyster, 2005, p. 21). Students diagnosed with dyslexia in Norway are typically assigned
18
19 lessons with a specially trained teacher each week. These are typically reading-focussed but
20
21 can also involve spelling practice and occasionally written composition. Students diagnosed
22
23 with dyslexia are typically provided with a personal computer and assistive software to
24
25 support spelling and text planning. From 8th grade all examinations and most other major
26
27 assignments in Norwegian schools are usually written on computers. From the start of upper-
28
29 secondary school all students are required to have their own laptop.
30
31
32
33

34 35 **Writing tasks**

36
37 Students wrote two expository texts, with order and topic counterbalanced across the normal /
38
39 masked-text condition. Task statements were “*Bør lekser avskaffes? Skriv en*
40
41 *sammenhengende tekst der du argumenterer for og mot dette.*” (“Should teachers cease to
42
43 give students homework? Write an essay arguing the pros and cons of giving students
44
45 homework.”) and “*Bør Ungdomskortet være gratis? Skriv en sammenhengende tekst der du*
46
47 *argumenterer for og mot dette.*” (“Should public transportation for young people be free?
48
49 Write an essay arguing the pros and cons of having free public transportations for young
50
51 people”). Students wrote under “examination conditions” in a quiet room in their own school.
52
53 They were tested in matched pairs. This provided some control over task duration without
54
55
56
57
58
59
60

1
2
3
4 imposing a task-duration expectation across the whole sample (with negative consequences
5 for the validity of the process data). Students were asked to approach the task as they would a
6 normal in-class writing assignment. They were told that the person testing them, together
7 with another researcher, would read their text. They were encouraged to do their best. Before
8 writing in the masked text condition students were assured that writing in this condition is
9 possible and what they wrote, although not visible, was stored.

18 **Measures**

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

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

33 **Text-based assessment of final texts.** We recorded, for final texts, text length (number
34 of word tokens), number of words incorrectly spelled, type-token ratio (as a measure of
35 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:

⁹⁴⁵T¹³⁰h¹¹²e¹¹⁷ _ ²⁴⁹b¹¹⁹o¹⁰²y⁹⁵ _ ³⁴⁰s¹²⁰w¹⁴⁰a¹⁰²m¹¹⁰.

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]

1
2
3
4 Ratings of the quality of students' spelling-corrected texts were significantly lower for
5
6 weak decoders on all quality dimensions (Main effects of group: organization, $F(1,50) = 21$,
7
8 $p < .001$, $\eta^2 = .30$; theme development $F(1,50) = 18$, $p < .001$, $\eta^2 = .27$; vocabulary $F(1,50) =$
9
10 26 , $p < .001$, $\eta^2 = .36$) . In weak decoders only, masking the text resulted in poorer
11
12 organization, although this effect was small ($F(1,50) = 4.8$, $p = .033$, $\eta^2 = .08$ for the group-
13
14 by-condition interaction). We found no other statistically significant effects.

15
16
17 Weak decoders produced shorter texts than controls ($F(1,50) = 5.4$, $p = .024$, $\eta^2 = .10$ for
18
19 the main effect of group). We found no effects involving condition.

20
21
22 Statistical analysis for the remaining product measures statistically controlled for text
23
24 length, as detailed above. Number of spelling errors and word frequency were log
25
26 transformed prior to analysis. As predicted, weak decoders made many more spelling errors
27
28 than controls ($F(1,47) = 51$, $p < .001$, $\eta^2 = .47$ for the main effect of group). As might be
29
30 expected, there were also reliably more spelling errors in the masked condition ($F(1,47) = 25$,
31
32 $p < .001$, $\eta^2 = .32$ for the main effect of condition). The group-by-condition interaction was
33
34 not significant. Weak decoders produced text with a lower ratio of open-class to closed class
35
36 words, suggesting less informationally-rich text, although this effect was small ($F(1,47) =$
37
38 8.3 , $p = .006$, $\eta^2 = .07$ for the main effect of group). Weak decoders also tended to use
39
40 slightly shorter open-class (content) words ($F(1,47) = 11.6$, $p = .001$, $\eta^2 = .30$ for the main
41
42 effect of group). We found no effects involving condition, and no significant effects for word
43
44 frequency or type-token ratio.
45
46
47
48

49 **Writing processes**

50
51 In reporting process measures we first report total time on task. We then examine differences
52
53 in key-press latency associated with location within the text (start of a sentence, start of a
54
55 word, within a word, end of word). We then focus on just word-level effects, looking first at
56
57
58
59
60

1
2
3
4 the probability of within-word editing, and finally at effects of spelling accuracy on key-press
5
6 latency.
7

8
9 With the exception of analysis of total time on task, data comprised key-presses, nested
10 within writing task, nested within student. Statistical inference involved testing linear mixed
11 effects models, starting with a zero (intercept-only) model with random by-subject intercepts
12 and random by-subject slopes for writing condition (masked vs. normal). We used χ^2 tests
13 based on change in log-likelihood to determine whether differences in model fit were
14 statistically significant. Statistical significance of parameters was established by evaluating
15 against the z distribution.
16
17
18
19
20
21
22

23
24 **Total time on task.** Time on task is reported in Table 2. We found no statistically
25 significant effect of group. We did, however, find an main effect of condition: Masking the
26 text resulted in slightly less time-on-task in both groups ($F(1,50) = 7.1, p = .01, \eta^2 = .11$ for
27 the main effect of condition in a group-by-condition ANOVA).
28
29
30
31

32 [insert Table 2 near here]
33
34

35 **Key-press latencies.** Key-press latencies were strongly negatively skewed. This skew
36 exists in all response time data. However in the present context atypically long latencies
37 cannot necessarily be dismissed as outliers: There is potential for a small number of long
38 pauses to result from, or cause, substantial disruption to ongoing production. We therefore
39 handle skew in two ways. In the first analysis that we report we trimmed to remove keystroke
40 latencies longer than 8 s and then log-transformed prior to inferential testing³. In a second
41
42
43
44
45
46
47
48
49

50
51 ³ We checked the effects of removing latencies longer than 8 s by repeating the analysis on
52 untrimmed data. There was no difference in pattern of effects, except in the sentence initial
53 position which showed by-group differences in the untrimmed data. There were,
54
55
56
57
58
59
60

1
2
3
4 analysis, for word-level pauses only, we explored the frequency of pauses in several latency
5
6 (pause duration) bins.
7

8
9 Table 3 gives keystroke latencies at six different locations within the text, by group and
10 by condition (masked-text vs. normal). We tested four nested models, starting with an
11 intercept-only model with random by-participant intercepts, and random by-participant slopes
12 for both location and condition. Each successive model gave better fit (respectively, $\chi^2(7) >$
13 $100, p < .001$; $\chi^2(11) > 100, p < .001$; $\chi^2(5) = 12.1, p = .03$). We followed this analysis with
14 three simultaneous linear contrasts conducted separately at each level of location (Table 3,
15 lower panel).
16
17
18
19
20
21
22
23

24 [insert Table 3 near here]
25

26 Main findings were as follows: Within-word latencies were on average 27% (47 *ms*)
27 longer for weak decoders. Masking the text slowed within-word latencies slightly for both
28 groups. Pre-word latencies, which we assume are associated primarily with word-level
29 planning, were 36% (115 *ms*) longer for weak decoders, consistent with the word-level-focus
30 hypothesis. There was no effect of masking. Pre-sentence latencies – associated additionally
31 with planning content and syntax – were substantially longer for weak decoders than for
32 controls, and longer in both groups when the text was masked.
33
34
35
36
37
38
39
40
41

42 In a second analysis, we explored word-level effects by categorizing within-word and
43 pre-word latencies into pause-length bins (.5 to 1 *s*, 1 to 2 *s*, 2 to 3 *s*, and > 3 *s*). We used
44 logistic mixed-effects regression methods to predict the probability of bin membership on the
45 basis of group and condition. We tested separate models at each pause-duration bin, with the
46 binomial is-in-bin vs. is-not-in-bin as dependent variable. We fitted full factorial models with
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

1
2
3
4 main effects for group (weak decoders vs. control) and condition (normal vs. masked), and
5
6 their interaction, and random by-subject intercepts and condition slopes. Comparing fit of this
7
8 model with that of the zero (intercept and random effects) model gave $\chi^2(3) = 9.2, p = .027$
9
10 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;
11
12 and $\chi^2(3) > 13, p < .005$ in other cases except mid-word > 3 s pauses and pre-word 2 to 3 s
13
14 pauses, neither of which gave significantly better fit.
15

16
17 Parameter estimates are reported in Table 4. Findings were as follows: Within word
18
19 pausing in both .5 to 1 s and 1 to 2 s ranges was 2.6 times more probable in weak decoders
20
21 than in controls. Pauses longer than 2 s were very rare in both groups. Pre-word pauses longer
22
23 than 1 s were approximately twice as common in weak decoders, although this effect failed to
24
25 reach significance for pauses in the 2 to 3 s range. Around 1 in 12 (non-sentence-initial)
26
27 words written by weak decoders in the normal condition were preceded by pauses for longer
28
29 than 2 s, compared to 1 in 26 for controls. Masking had no clear effect.
30
31

32
33 **Relationship between word-level latencies and spelling accuracy.** Here we focus just
34
35 on within-word and word-initial latencies for words that were produced without internal
36
37 editing. We determined the relationship between these latencies and the spelling accuracy of
38
39 the resulting word. This addresses questions about whether writers who fail to spell a specific
40
41 word correctly hesitate more prior to and/or during its production. We analyzed just open-
42
43 class words to reduce confound with grammatical function. We also statistically controlled
44
45 for word length and frequency. We modeled within-word and word-initial latencies
46
47 separately, starting with a zero model, then adding word length and frequency as covariates,
48
49 then main-effects of group, condition, and whether or not the word was correctly spelled,
50
51 then two way interactions between these factors, and finally a full factorial model. For both
52
53 within-word and before-word latencies the covariates-only model gave better fit than the zero
54
55
56
57
58
59
60

1
2
3
4 model⁴ ($\chi^2(2) > 100, p < .001$ for both within and before words). Adding main effects further
5 improved fit ($\chi^2(3) = 103, p < .001$ within words; $\chi^2(3) = 36, p < .001$ before words), with
6 significant effects for condition and group – repeating findings from the previous section –
7 and a significant effect of whether or not the word was correctly spelled ($z = 8.3, p < .001$
8 within words; $z = 5.0, p < .001$ before words). We followed this with simultaneous pairwise
9 contrasts, within groups, comparing correctly and incorrectly spelled words.
10
11
12
13
14
15
16

17 Estimated means and contrasts are given in Table 5. Findings were clear, consistent, and
18 unexpected. In both groups words that were spelled incorrectly were associated with *shorter*
19 preparation time before output was initiated and more rapid production when the word was
20 being typed.
21
22
23
24
25

26 **Word-level editing.** We identified words as having been “internally edited” if at any
27 point between the initial keypress and the keypress for the word-terminal character (space or
28 punctuation) there were any delete or horizontal cursor-key presses⁵ (henceforth *edit key-*
29 *presses*). We tested logistic mixed-effects regression models, predicting the probability that a
30 word was produced with internal editing on the basis of group and condition, first adding
31 main effects to an intercept-only (zero) model, and then adding the interaction. These models
32 showed significantly better fit (respectively, $\chi^2(2) = 34, p < .001$ and $\chi^2(1) = 4.7, p = .03$). As
33 can be seen from Table 2, weak decoders in the seen-text condition were slightly less likely
34 to edit words. This effect was not present when the text was masked.
35
36
37
38
39
40
41
42
43
44
45

46 ⁴ Latencies were longer for longer words and lower frequency words, both before and within
47 words. $z > 6, p < .001$ in all cases.
48
49

50 ⁵ All cursor moves were by cursor key and all deletes by backspace. Vertical cursor-key
51 presses, although potentially associated with editing, were omitted because this was also the
52 mechanism by which text was vertically scrolled.
53
54
55
56
57

1
2
3
4 We then explored effects associated just with words that *were* edited during production.
5
6 We found no evidence in the extent of editing across group and condition, measured in terms
7
8 of edit keypresses. Spelling accuracy for edited words was slightly lower for weak-decoders
9
10 (Table 2; $\chi^2(2) = 11.5, p = .003$ for the main effects model, as detailed above, with $z = 3.47, p$
11
12 $= .001$ for the effect of group), but there were no main effects of condition or group-by-
13
14 condition interaction. In response to a reviewer's suggestion, we also explored the possibility
15
16 that, in the normal condition, editing took longer per keystroke for weak decoders, relative to
17
18 what would be expected based on their normal (non-edited word) within-word production
19
20 rate. We did not find evidence that this was the case.
21
22

23
24 Overall, we found no evidence that weak decoders were more likely to make changes to
25
26 their text, based on the ratio of the number of edit keypresses to the number of character-key
27
28 presses (Table 2). As might be expected, within-word editing was less common when the text
29
30 was masked ($F(1,102) = 40, p < .001$ for a group x condition ANOVA). This effect was
31
32 independent of group.
33
34

35 36 Discussion

37
38 Our study sampled students in mainstream upper-secondary education writing in a language
39
40 with relatively shallow orthography. Within this specific context we aimed to answer three
41
42 questions: 1. Do students who struggle with decoding produce poorer quality texts? 2. Is there
43
44 evidence of a word-level focus in weak-decoder's writing processes? 3. If so, is this primarily
45
46 a production deficit or does it result from problems reading what they are or have just written
47
48 (the "monitoring hypothesis")? We will address each question in turn.
49
50

51
52 Weak decoders in the normal writing condition made, on average, more than four times
53
54 as many spelling errors as controls. This is consistent with previous findings (Connelly et al.,
55
56 2006; Sterling et al., 1998; Tops et al., 2013; Wengelin, 2002). Spelling aside, weak decoders

1
2
3
4 produced text that was rated as having poorer use of vocabulary, poorer theme development,
5 and being less well structured. Evidence for quality differences in this study was stronger
6 than in the only previous study that we are aware of that has tested a broadly comparable
7 sample and used a similar writing task (Connelly et al., 2006). Our findings are consistent
8 with the hypothesis that spelling difficulties hamper students' ability to produce higher-level
9 text features, although, as we suggest below, correlation does not necessarily imply causality.
10
11 Third-factor explanations are equally plausible.

12
13 Findings from analysis of key-press latencies provided partial evidence for the word-
14 level-focus hypothesis. Latencies were longer for weak decoders, relative to controls, at all
15 text locations. Word-initial latencies, which are likely to be associated with retrieving
16 spelling, showed clear effects both in mean duration and, perhaps more importantly, in the
17 proportion of longer pauses. This suggests that weak decoders found word-level planning
18 generally more effortful and experienced a larger number of interruptions to fluent
19 production.

20
21 Word-level hesitation appeared, however, to be entirely associated with production
22 rather than reading. According to the "monitoring hypothesis" disfluency results in part from
23 students with dyslexia struggling to read the word that they are writing or have just written
24 and/or identifying this word as misspelled. If weak decoders in our sample had spent time
25 looking back at words they had just written and puzzling over whether or not they were
26 correctly spelled, then we would expect differences between groups to disappear when they
27 were prevented from seeing what they had written. This was not what we found. In fact in
28 both groups word-end latencies were typically similar to or slightly shorter than within-word
29 latencies. This suggests that the move from the word-terminal character to the space bar was
30 not associated with any additional processing beyond that which occurred within-word. It is
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 also possible that inflated word-initial latencies in weak decoders occurred because students
5
6 looked back at the just-produced word (or earlier words). However if this were the case then
7
8 the difference between groups should disappear in the masked-text condition. Again, this is
9
10 not what we found. More generally, in the context of the relatively short writing task used in
11
12 this study, we found that masking the text, and therefore preventing reading what was
13
14 written, did not have any negative consequences for text quality in either group. This result is
15
16 perhaps counterintuitive, but it is consistent with previous findings (Gould & Boise, 1978;
17
18 Hull & Smith, 1983; but see Olive & Piolat, 2002).
19
20

21
22 Compared to controls, weak decoders were somewhat slower in typing words, after the
23
24 initial character. They were, however, no more likely than controls to engage in within-word
25
26 editing and there was no evidence that when they did edit they did so more extensively. The
27
28 writing behavior of weak decoders in our sample therefore was not consistent with anecdotal
29
30 accounts of very extensive and repeated within-word delay and editing by writers with
31
32 dyslexia. Failure to retrieve an accurate spelling was, in fact, associated with *shorter*
33
34 latencies, in both groups, both before and within words. Thus, although weak decoders found
35
36 retrieval of correct spelling more effortful relative to controls, there was no evidence that
37
38 failure to retrieve correct spelling typically introduced further disfluency. On the contrary,
39
40 production tended to be faster when spelling was incorrect.
41
42
43

44 Our results suggest, therefore, that weak decoders (and, by extension, students with a
45
46 dyslexia diagnosis) tend to make more spelling errors and produce text that was judged as
47
48 lexically, structurally, and thematically weaker than that of peers with average decoding
49
50 ability, even after spelling was corrected. Our results also suggest that, for spelling rather
51
52 than decoding reasons, weak decoders take longer, and are more likely to hesitate, when
53
54 preparing words. What is less clear is whether these effects are causally related. Is it the case,
55
56
57
58
59
60

1
2
3
4 as we hypothesized in our introduction, that poorer text in weak decoders was caused by
5
6 diversion of attention to the word level reducing resources available for the higher-level
7
8 processes necessary to create well-formed text? Some evidence in support of this argument
9
10 perhaps comes from the fact that we found substantially longer sentence-initial latencies in
11
12 weak decoders. In experiments exploring written sentence production, Nottbusch (2010)
13
14 found sentence-initial latencies of around 1300 *ms* (timed between subjects first seeing a
15
16 stimulus array and the first keystroke of their response). This value is only slightly higher
17
18 than the sentence-initial latencies found for weak-decoders in this study, but much higher
19
20 than for controls. It may be that control students were able to start planning the content and
21
22 syntax of their next sentence in parallel with typing. The higher demands of word-level
23
24 processing for weak-decoders mean that in these students sentence planning was partly or
25
26 entirely delayed until the inter-sentence pause.
27
28
29

30
31 However, even if word-level focus reduces parallelism in production, it is not clear that
32
33 this will necessarily have consequences for higher-level text features. It is equally plausible
34
35 that relatively weak structure and content are due to effects of reading deficit on prior
36
37 learning (see Stanovich, 1994, for an analogous argument concerning the relationship
38
39 between reading and verbal ability). The primary deficit associated with poor decoding skills
40
41 (and therefore a probable dyslexia diagnosis) is difficulty with reading. This results in both a
42
43 poorer comprehension (Ransby & Swanson, 2003) and a general tendency to read less (Mol
44
45 & Bus, 2011). During their school careers students with dyslexia will, therefore, be less well
46
47 placed to learn, through reading, the linguistic strategies and conventions that will support
48
49 their own production of coherent text. This, rather than struggling with spelling when writing,
50
51 may be a more parsimonious explanation for weak decoders' relatively poor text.
52
53
54
55
56
57
58
59
60

1
2
3
4 In conclusion, therefore, our findings suggest that struggling with decoding has negative
5 consequences for students' written products and has measurable effects on writing processes.
6 We did not, however, find evidence of the kinds of extensive within-word editing or frequent
7 very long pauses that would be strong evidence that poorer quality text results directly from a
8 word-level focus. Whether or not stronger effects would emerge with, for example, a younger
9 sample writing in a less transparent orthography is a question that could usefully be explored
10 in future research.
11
12
13
14
15
16
17
18
19

20 Acknowledgements

21 References

- 22
23
24
25
26 Bereiter, C., Burtis, P., & Scardamalia, M. (1988). Cognitive operations in constructing main
27 points in written composition. *Journal of Memory and Language*, 27(3), 261–278.
28 doi:10.1016/0749-596X(88)90054-X
29
30
31
32
33 Berninger, V., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., ...
34
35 Graham, S. (2002). Teaching spelling and composition alone and together: Implications
36 for the simple view of writing. *Journal of Educational Psychology*. doi:10.1037/0022-
37 0663.94.2.291
38
39
40
41
42 Berninger, V. W. (2000). Dyslexia the invisible, treatable disorder: The story of Einstein's
43 ninja turtles. *Learning Disability Quarterly*, 23(3), 175–195. doi:10.2307/1511163
44
45
46 Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing
47 problems in developmental dyslexia: under-recognized and under-treated. *Journal of*
48 *School Psychology*, 46(1), 1–21. doi:10.1016/j.jsp.2006.11.008
49
50
51
52
53 Bertram, R., Tønnessen, F. E., Strömqvist, S., Hyönä, J., & Niemi, P. (2015). Cascaded
54 processing in written compound word production. *Frontiers in Human Neuroscience*,
55
56

- 1
2
3
4 9(April), 1–10. doi:10.3389/fnhum.2015.00207
5
6 Bonin, P., Malardier, N., Méot, A., Fayol, M., & Meot, A. (2006). The scope of advance
7
8 planning in written picture naming. *Language and Cognitive Processes*, 21(1-3), 205–
9
10 237. doi:10.1080/01690960400002141
11
12 Connelly, V., Campbell, S., MacLean, M., & Barnes, J. (2006). Contribution of Lower Order
13
14 Skills to the Written Composition of College Students With and Without Dyslexia.
15
16 *Developmental Neuropsychology*, 29(1), 175–196. doi:10.1207/s15326942dn2901_9
17
18 Damian, M., & Stadthagen-Gonzalez, H. (2009). Advance planning of form properties in the
19
20 written production of single and multiple words. *Language and Cognitive Processes*,
21
22 24(4), 555–579. Retrieved from <Go to ISI>://000265359600004
23
24
25
26 Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling
27
28 regularity affects both writing latencies and durations. *Journal of Experimental*
29
30 *Psychology: Learning, Memory, and Cognition*, 32(6), 1330–1340. doi:10.1037/0278-
31
32 7393.32.6.1330
33
34
35 Fayol, M. (1999). From on-line management problems to strategies in written composition. In
36
37 M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing: processing capacity*
38
39 *and working memory effects in text production* (pp. 15–23). Amsterdam: Amsterdam
40
41 University Press.
42
43
44 Fayol, M., Largy, P., & Lemaire, P. (1994). A special case of agreement error in French: The
45
46 confusion of noun/verb inflections. In G. Rijlaarsdam, H. van den Burgh, & M. Couzijn
47
48 (Eds.), *Theories, models and methodology in writing research* (pp. 169–212).
49
50 Amsterdam: Amsterdam University Press.
51
52
53 Flower, L., & Hayes, J. (1980). The dynamics of composing: Making plans and juggling
54
55 constraints. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp.
56
57
58
59
60

- 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.

- 1
2
3
4 Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of*
5
6 *Dyslexia*. doi:10.1007/s11881-003-0001-9
7
- 8
9 Maggio, S., Lété, B., Chenu, F., Jisa, H., & Fayol, M. (2012). Tracking the mind during
10
11 writing: Immediacy, delayed, and anticipatory effects on pauses and writing rate.
12
13 *Reading and Writing*, 25, 2131–2151. doi:10.1007/s11145-011-9348-1
14
- 15
16 McCutchen, D. (1996, September). A capacity theory of writing: Working memory in
17
18 composition. *Educational Psychology Review*. doi:10.1007/BF01464076
19
- 20
21 Miller-Guron, L. (1999). *Wordchains: A word reading test for all ages*. Windsor, England:
22
23 NFER-Nelson.
- 24
25 Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure
26
27 from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267–296.
28
29 doi:10.1037/a0021890
- 30
31 Nottbusch, G. (2009). Grammatical planning, execution, and control in written sentence
32
33 production. *Reading and Writing*, 23(7), 777–801. doi:10.1007/s11145-009-9188-4
34
- 35
36 Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of
37
38 research on writing processes coordination. *Journal of Writing Research*, 6(2), 173–194.
39
40 doi:10.17239/jowr-2014.06.02.4
41
- 42
43 Olive, T., & Piolat, A. (2002). Suppressing visual feedback in written composition: Effects
44
45 on processing demands and coordination of the writing processes. *International Journal*
46
47 *of Psychology*, 37(4), 209–218. doi:10.1080/00207590244000089
- 48
49 Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual*
50
51 *Review of Psychology*, 52, 629–651. doi:10.1146/annurev.psych.52.1.629
52
- 53
54 Ransby, M. J., & Swanson, H. L. (2003). Reading comprehension skills of young adults with
55
56 childhood diagnoses of dyslexia. *Journal of Learning Disabilities*, 36(6), 538–555.
57
58
59
60

1
2
3
4 doi:10.1177/00222194030360060501

5
6 Stanovich, K. S. (1994). Annotation: Does Dyslexia Exist. *Journal of Child Psychology and*
7
8 *Psychiatry*, 35(4), 579–595. doi:10.1111/j.1469-7610.1994.tb01208.x

9
10 Sterling, C., Farmer, M., Riddick, B., Morgan, S., & Matthews, C. (1998). Adult Dyslexic
11
12 Writing. *Dyslexia*, 4, 1–15. doi:D-DYS87>3.0.CO;2-F

13
14 Sumner, E., Connelly, V., & Barnett, A. L. (2013). Children with dyslexia are slow writers
15
16 because they pause more often and not because they are slow at handwriting execution.
17
18 *Reading and Writing*, 26(6), 991–1008. doi:10.1007/s11145-012-9403-6

19
20 Sumner, E., Connelly, V., & Barnett, A. L. (2014). The influence of spelling ability on
21
22 handwriting production: children with and without dyslexia. *Journal of Experimental*
23
24 *Psychology. Learning, Memory, and Cognition*, 40(5), 1441–7. doi:10.1037/a0035785

25
26 Tops, W., Callens, C., van Cauwenberghe, E., Adriaens, J., & Brysbaert, M. (2013). Beyond
27
28 spelling: The writing skills of students with dyslexia in higher education. *Reading and*
29
30 *Writing*, 26(5), 705–720. doi:10.1007/s11145-012-9387-2

31
32 Torrance, M. (2015). Understanding Planning in Text Production. In C. MacArthur, S.
33
34 Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (2nd Editio.). New
35
36 York: Guildford Press.

37
38 Totereau, C., Thevenin, M., & Fayol, M. (1997). The development of the understanding of
39
40 number morphology in written French. In C. Perfetti, L. Rieben, & M. Fayol (Eds.),
41
42 *Learning to Spell* (pp. 97–114). Hillsdale, NJ: Erlbaum.

43
44 Wechsler, D. (2006). *Wechsler Individual Achievement Test - Second UK Edition (WIAT-II*
45
46 *UK)*. Oxford: Pearson Assessment.

47
48 Wengelin, Å. (2002). *Text production in adults with reading and writing difficulties*
49
50 *(Gothenburg Monographs of Linguistics, 20)*. Gothenburg: Gothenburg University.

1
2
3
4 Wengelin, Å. (2007). The word level focus in text production by adults with reading and
5
6 writing difficulties. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and*
7
8 *Cognition: Research and Applications* (pp. 68–82). Amsterdam: Elsevier.
9

10 Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system.
11
12 *Applied Psycholinguistics*. doi:10.1017/S0142716400010122
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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, * $p < .05$, ** $p < .005$; ^aGroup 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 <i>b^o^y</i>	Word-end <i>The^ boy^</i>	Pre-word <i>The ^boy</i>	Pre-sentence <i>swam. ^Then</i>
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, <i>ns</i>	2.4, .019
Control: Normal vs. Masked	4.6, < .001	3.2, .003	.20, <i>ns</i>	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, <i>ns</i>	1.7, <i>ns</i>
condition	1.7, <i>ns</i>	.81, <i>ns</i>	1.5, <i>ns</i>	1.8, <i>ns</i>
interaction	2.4, .02	1.7, <i>ns</i>	2.3, <i>ns</i>	1.4, <i>ns</i>
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, <i>ns</i>	2.5, .011	1.7, <i>ns</i>	4.0, < .001
condition	2.3, .019	.01, <i>ns</i>	.06, <i>ns</i>	1.7, <i>ns</i>
interaction	.85, <i>ns</i>	.25, <i>ns</i>	.12, <i>ns</i>	1.6, <i>ns</i>

Note. ^aEffects (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.