STRUGGLING WRITERS IN ELEMENTARY SCHOOL: CAPTURING DRIVERS

OF PERFORMANCE

Julie E. Dockrell

Department of Psychology and Human Development

UCL Institute of Education

20 Bedford Way

London WC1H 0AL

Vincent Connelly

Department of Psychology

Oxford Brookes University

Gipsy Lane,

Oxford OX3 0BP

Barbara Arfè

Department of Developmental Psychology and Socialisation

University of Padova

Via Venezia

35131 Padova

Corresponding Author: Julie Dockrell j.dockrell@ucl.ac.uk

Abstract

Conceptualising the difficulties experienced by struggling writers in middle elementary school is of both theoretical and practical importance. To further our understanding of the problems experienced by struggling writers we aimed to identify the writing measure which best discriminated struggling writers from their peers, and the proximal and distal factors which which contributed to performance. The performance of 96 students (Mean age 0;4), 39 of whom were independently identified as struggling writers using a norm referenced standardised test, was examined at word, sentence, and text level. Standardised measures of transcription, oral language and working memory were collected. The extent to which independent product and process writing measures accurately identified the students was tested using ROC analyses. The skills which underpinned performance were examined using regression analyses and path analysis.

Written sentence generation was the most sensitive and specific measure to identify struggling writers at this point in development and, was concurrently predicted by both oral sentence level skills, handwriting fluency and listening span. Path models demonstrated that oral language contributed both directly and indirectly to sentence level writing. Implications for developmental models of writing and support for struggling writers are discussed.

A significant number of elementary school students struggle with writing. Struggling writers (SWs) experience difficulties with a broad range of the component skills which underpin writing. This is challenging in terms of the conceptualisation of writing development, understanding the relative impact of different skills (Abbott, Berninger, & Fayol, 2010; Kent & Wanzek, 2016; Kim & Schatschneider, 2017) and, as such, targeting key skills to support writing development (Olinghouse & Santangelo, 2010). Current writing models are informative about the predictors of writing skills, but they lack specificity about the relationships between text quality and productivity, and those highly related skills such as spelling, handwriting fluency and the wider measures of working memory (WM), oral language, and reading (Kim & Schatschneider, 2017). This lack of specificity limits progress in developing writing models and, in turn, means that the the relationships between the component skills of writing across development are ill defined. Moroever, interpretation of these relationships is often limited because of the focus on single measures or skills (Kent & Wanzek, 2016) or single developmental disorders (Connelly, Dockrell, & Barnett, 2012).

To address these limitations in our understanding of the difficulties experienced by SWs the current study aims to 1) identify specific and sensitive discriminators of writing performance in SWs and 2) capture the direct and indirect factors that drive performance on these discriminators of writing. To address these aims the performance of SWs is compared to their age matched typical writers (TWs) on measures of written text production and on measures associated with the production of written text. Given the extant literature, outlined below, we reasoned that the quality of writing and a writing process measure would distinguish SWs from TWs in middle elementary school and that performance would be, primarily, underpinned by transcription skills (handwriting and spelling), but anticipated that both oral language and WM would contribute to writing performance.

1.1. Elementary school writing

Writing products and writing processes in middle elementary school are influenced by a number of cognitive skills, either directly or indirectly (MacArthur & Graham, 2016). In the initial phases of learning to write in English, transcription skills, often measured by fluency in handwriting letters and spelling accuracy, account for the majority of variance in written composition (Berninger et al., 1992; Graham et al., 1997). Difficulties in transcription interfere with both the amount and quality of writing (Graham, et al., 1997); where a lack of automaticity in the retrieval of accurate spellings and/or the production of letters on the page can inhibit the quality and fluency of written expression (Alves & Limpo, 2015). The variance explained by transcription skills reduces as developing writers progress through elementary school (Berninger & Swanson, 1994), but continues to constrain the performance of poor writers into adolescence (Connelly et al., 2012).

1.2 Struggling writers

The difficulties experienced by SWs are evident in terms of their writing products at word, sentence and text level (Graham & Harris, 2005), and the lack of fluency with which they produce written texts (Garcia, Crespo, & Bermudez, 2017). SWs have difficulties with transcription, construct short, choppy, or incomplete sentences with numerous errors in spelling, punctuation and capitalisation. Together these difficulties result in texts that are of significantly poorer quality than their typically developing peers. Of importance for both theory and practice is the extent to which the dominance of these aspects may vary over time, within children and which factors underpin performance. Identifying relevant skills in middle elementary school, when transcription skills are more secure, could lead to the elaboration of theoretical frameworks for the development of writing, and more reliable and valid age appropriate writing assessments.

1.3 Proximal and distal factors influencing text writing

A number of frameworks/models exist to conceptualise the development of writing which focus on the interacting components necessary for writing (Berninger & Winn, 2006: Kim & Schatschneider, 2016) or other factors such as, working memory (WM) capacity limitations (McCutchen 2012), or the writing process (Chenoweth and Hayes, 2001). An alternative approach is to draw an analogy with the literature on reading development and to distinguish the proximal and distal factors that support writing development. From this perspective, proximal factors can be described as the direct building blocks in the production of written text, whereas distal factors reflect the underpinning competencies that support writing, and thus play a more indirect role (Zoccolotti, De Luca, Marinelli, & Spinelli, 2014). Transcription skills that directly impact on the production of written text can be conceptualised as proximal factors. By contrast, competencies which work more indirectly and mediate the production of written text, can be considered as distal factors. Research has typically focused on the proximal causes, or the links between them, in the identification of SWs (see Connelly, et al., 2012 for a review). However, as children master these proximal skills then distal factors are hypothesised to play a greater role in text generation, even though the variance accounted for in the children's written products is often small (Kent & Wanzek, 2016; Kim & Schatschneider, 2017). Here we consider these more distal underlying competencies to include oral language, reading and WM.

To generate text, children need to access cognitive representations and this involves fluent use of different levels of language to translate these into written products (Abbott et al., 2010; Chenoweth & Hayes, 2001). Oral language skills account for independent variance in children's written text products, although the precise linguistic driver is a matter of debate (Mackie, Dockrell, & Lindsay, 2013). Vocabulary, oral 'sentence[1]' skills and discourse level language, are all associated with writing development (Abbott & Berninger, 1993; Arfé &

---

[1] Scare quotes are used to denote oral 'sentence' skills as we actually speak in utterances.

Pizzocaro, 2016; Arfé, Dockrell, & De Bernardi, 2016; Kim et al, 2014; McCutchen, Covill, Hoyne, & Mildes, 1994; Olinghouse & Leaird, 2009). Isolating their effects depends on the age of the children assessed, the way in which writing is assessed and whether the children assessed have additional learning needs (Connelly et al., 2012; Dockrell, Lindsay, & Connelly, 2009). Current studies typically report only weak to moderate correlations between measures of oral language and the quality of student's writing, and weak correlations with productivity (Apel & Apel, 2011; Kent & Wanzek, 2016; Olinghouse & Graham, 2009; Olinghouse & Leaird, 2009), suggesting that while oral language is important it is not as significant or robust as the impact of proximal factors. Consequently, oral language may best be conceptualised as a distal factor supporting the development of written text production. Distal however, does not mean insignificant as there is preliminary evidence that oral 'sentence' fluency supports written text generation over time and across languages (Savage, Kozakewich, Genesee, Erdos, & Haigh, 2017).

Reading also influences written text production. Word reading is associated with transcription skills, and word recognition skills consistently predict spelling abilities at all elementary grade levels (Abbott & Berninger, 1993). Additionally, improvement in word reading leads to an improvement in spelling (Ahmed, Wagner, & Lopez, 2014). By corollary, poor reading comprehension impacts on text level writing, where children with poorer reading comprehension, but age appropriate spelling, produce texts which are more limited and less sophisticated in comparison to age matched peers (Cragg & Nation, 2006). Bidirectional relations between reading and writing exist (Abbott et al., 2010), but recent evidence suggests that reading-to-writing conceptualisations are superior, especially for word and text levels of writing (Ahmed et al., 2014). Again, only moderate associations between writing and measures of reading are reported (Kent & Wanzek, 2016).

The coordination of the components of text production place heavy demands on WM resources (McCutchen, 1996; Olive, Favart, Beauvais, & Beauvais, 2009). If a child needs to search for the letters to create a word, then ideas and plans can be lost as they overload the capacity of WM (Graham et al., 1997). WM is also an independent longitudinal predictor of later writing disabilities in elementary school (Costa, Green, Sideris & Hooper, 2017) and listening span, a measure of WM, a significant predictor of text generation in upper elementary school (Swanson 1996).

In sum, based on the current research evidence for children in the middle years of elementary school, both handwriting fluency and spelling serve as the direct building blocks of the production of written text and can be conceptualised as proximal factors in writing development. Written text production is further supported by more distal factors, although their contribution to the children's writing is harder to detect and the variance detected in children's writing products is often small (but see Kim et al., 2017).

1.4 Assessing writing

By middle elementary school children are expected to produce multi-paragraph texts and so the amount of written product and, in theory, the quality of writing that children produce often rises sharply. Measuring writing performance can become more challenging and this impacts on the ability to identify SWs. The reliability of analytic scales are reportedly too low to guide decision making and planning (Graham, Hebert & Harris, 2011; He, Anwyll, Glanville & Deavall, 2013; Williams, Larkin, & Blaggan, 2013). Moreover, when only one dimension, such as text quality, is examined, our understanding of writing development is limited, hence the need to consider productivity and accuracy (Kim, Al Otaiba, & Wanzek, 2015). Productivity is often examined at the word level and calculated by counting the number of legible words produced. By contrast, measures of accuracy can capture both grammatical and spelling accuracy either in sentences or texts.

Recent research has highlighted that an important additional core measure of writing competence is the fluency with which single written sentences are generated (Arfé, et al., 2016). The ability to translate ideas into written sentences flexibly influences text production, and develops throughout elementary school (Arfé & Pizzocaro, 2016; Berninger, Nagy, & Beers, 2011) and is directly associated with dimensions of writing proficiency including written text productivity, accuracy and overall text quality (Arfé & Pizzocaro, 2016; Arfé et al., 2016). To express their thoughts in writing, writers translate and organize their ideas into morpho-syntactic structures, which correspond to the written sentences of the ongoing text. Written sentence generation measures tap the demands of text generation, assessing the child's ability to retrieve and use lexical and grammatical knowledge (e.g. Arfé et al., 2016; Berninger et al., 2011). Assessing sentences within texts is problematic as this confounds sentence generation skills with cognitive processes entailed in creating links with previous and subsequent ideas, placing greater demands on WM. Indicating that assessment of writing performance should include a measure of written sentence generation.

Another recent addition to our understanding of writing development has been the inclusion of measures of the writing process, typically measured by writing bursts. Writing bursts are defined by the numbers of words produced between writing pauses and are informative about the children's writing efficiency and were thought to derive from the translation processes in writers at work (Hayes, 2012). However, transcription processes have also been found to be important mediators of writing bursts. There is a yearly growth in burst length in children, with a corresponding reduction in pause length (Alves & Limpo, 2015). Moreover, text writing fluency, captured by writing bursts, accounts for a significant proportion of the variance in students' text quality (Kim, et al., 2017). Alves and Limpo (2015) argue that such measures may be of central importance in the identification of SWs. To our knowledge, this hypothesis has not been directly tested.

A comprehensive assessment of children's writing skills should then involve an evaluation of word, sentence and text production in combination with a writing process measure. Such an approach will allow a more detailed analysis of the challenges in written text production experienced by SWs.

1.5 The current study

The current literature demonstrates that SWs perform poorly across both writing product and process measures. However, evidence is lacking about which assessments of writing most reliably discriminate SWs from TWs in middle elementary school and which proximal and distal factors underpin performance on the discriminator task/s? We report the first study, to our knowledge, that 1) aims to capture the writing product and process assessment measures that best discriminate middle elementary SWs from TWs and 2) then examines the proximal and distal factors that predict poor performance on the discriminator variable.

In our study, children in mainstream middle elementary classes were identified as either a SW or a TW using a standardised norm referenced assessment of written text production. Performance in writing was then examined again on independent assessments of writing at word (words produced), sentence (Arfé et al., 2016), grammatical accuracy and text level quality (Keller-Margulis, Payan, Jaspers, & Brewton, 2016). Length of writing bursts was used to assess writing process efficiency during the writing task.

We reasoned that writing performance, at this point in development, could be sufficiently assessed at word, sentence and text level to allow us to identify measures which could reliably discriminate SWs from their peers. We predicted from the extant research literature that the SWs would perform significantly below their age-matched peers on all measures of writing performance that is productivity, accuracy, sentence generation and

writing bursts. Moreover, we reasoned from previous research discussed above that at this point in development both quality of text writing, a product measure, and burst length, a process measure would best discriminate the SWs from TWs.

ROC and discriminator analysis were used to test this hypothesis and establish which of the writing measures at word, sentence and text level most reliably distinguished SWs and TWs in middle elementary school. The ROC curve provides a diagnostic test evaluation where sensitivity (true rate) is plotted as a function of the false positive rate. The area under the ROC curve (AUC) is a measure of how well a parameter can distinguish between two diagnostic groups, in this case SWs and TWs and the most accurate discriminator can be confirmed by logistic regression.

We then examined which proximal and distal factors predicted the writing measure that had most accurately classified children as either a SW or a TW. Measures of proximal (handwriting fluency and spelling accuracy) and distal factors (oral vocabulary, oral 'sentence' generation, word reading, reading comprehension and listening span), were collected from the children during testing. Given the results in the published literature and the children's phase of writing development it was predicted that a) spelling would serve as the best proximal predictor and b) oral language, measured at the sentence level, would be the best distal predictor of performance on the discriminator variable (Savage et al., 2017).

2. Methods

2.1 Participants

Participants, between the ages of 8 and 11, were recruited from mainstream elementary schools as part of a larger study on writing development (Dockrell, Connelly, Walter & Critten, 2015) and screened to ensure that they were performing within the average range on a measure of non-verbal ability (Standard Scores (SS) > 85) and spoke English as

their first language. The final sample consisted of 96 pupils with a mean age of 111 months, ($SD = 10.9$).

Grouping of participants into typical writers and struggling writers

SWs scored below the 10th centile (SS < 80) in an independent standardised writing task. Thirty-nine (40%) were identified as SWs from their performance on the Wechsler Objective Language Dimensions (WOLD) standardised writing scale (Rust, 1996). Typical writers had standard scores within the average range and their performance differed significantly from the SWs with a large effect size, more than 2SD difference between the groups (TWs $M = 90.67$, $SD = 11.01$; SWs $M = 73.82$, $SD = 4.40$: $t(78.003) = 10.31$, $p < .0005$, $d = 2.15$). There was no significant difference in age between the two groups (SWs $M = 113$, $SD = 11.51$; TWs $M = 110.00$, $SD = 10.47$: $t(94) = 1.39$, $p = .17$), nor in the gender distribution of the groups ($X^2(1, N = 96) = 2.22$, $p = .14$). SWs scored within the average range for non-verbal ability but significantly lower than the TWs (SWs $M = 98.85$, $SD = 9.60$; TWs $M = 103.63$, $SD = 9.56$: $t(94) = 2.42$, $p = .02$, $d = .50$).

2.2 Measures

Written language tasks.

WOLD. The WOLD writing expression (Rust, 1996) was used to identify children's writing ability and to classify children as a SW or a TW. The child is asked to write a letter outlining their ideal house. Children have 15 minutes to complete the task. The written output can be scored holistically or analytically: reliability .89, correlation with Woodcock–Johnson Psycho-Educational Battery-Revised, Dictation = 0.72. The combined score on the analytic scale was used to assess the children's written text in relation to U.K standardised norms.

Curriculum based measure (CBM) of writing. This measure was used to assess writing productivity, accuracy and quality. Following standard CBM procedures for this age group (McMaster & Espin, 2007), students were asked to write for five minutes about their

'best day at school'. The prompt was presented in a typed format at the top of a page for writing and read to the students. The task was scored for productivity (total words written), grammatical accuracy (correct word sequences minus incorrect word sequences), and writing quality scale (Dockrell, et al., 2015). If the child only copied the prompt, they were classified as not producing written text and excluded from the study.

Written Sentence Generation. This task followed Arfé et al, (2016) and assessed children's fluency in generating ideas in single grammatically and semantically correct written sentences. Children received a sheet of lined paper with two word pairs and were asked to generate as many different sentences as they could using both words in five minutes. Before testing, the researcher demonstrated how to perform the task and participants practiced with a different word pair. A score of 1 was given to each sentence that was both grammatically and semantically correct. In scoring sentence accuracy errors in punctuation, capitalization or misspellings were not coded. Interrater reliability is good (94%) and test-retest reliability at a two-month interval is .62 (Arfé et al., 2016).

Proximal Factor Measures

Spelling. BAS-II (Elliott, Smith & McCulloch, 1997) spelling subtest. Students spell single words. The BAS-II provides norms for individuals aged 5 years to adult. The technical manual reports reliability ($r = .84$-$.96$), and test-retest reliability as $r = .64$ and validity $r = .63$.

Handwriting Fluency. Students write the alphabet in lower case in order, as many times as possible in 60 seconds as a measure of written handwriting fluency (DASH; Barnett, Henderson, Scheib, & Schulz, 2007). The DASH technical manual reports $\alpha = .83$-$.89$, indicating good internal consistency.

Distal Factor Measures

Language Measures

Expressive vocabulary. The CELF-4 UK (Semel, Wiig & Secord, 2006) expressive vocabulary task. Children produce the name for a pictured object, person or activity. Norms exist for individuals aged five to 21. Internal consistency is r = .80-.85 and reliability is .80.

Receptive vocabulary. British Picture Vocabulary Scale (BPVS III; Dunn & Dunn, 2009). Children hear a word and select a referent from four alternatives. The BPVS-III provides norms for individuals aged 3-16 years. Reliability is as 0.91 and validity with the WISC .76.

Oral 'sentence' generation. The formulated sentences subtest from CELF-4 UK (Semel et al., 2006) was used. Students formulate an oral response to an oral prompt with a stimulus picture. Internal consistency is r = .75-.89 and test-retest reliability r = .86.

Recalling sentences. The recalling sentences subtest from the CELF-4 UK (Semel et al., 2006) was used. Students imitate sentences provided by the examiner. The test manual reports r =.86 and validity as r = .76

Composite Language Measures. There were large and significant correlations found between the expressive and receptive vocabulary tasks (r (96) = .75, p < .001) and the oral 'sentence' generation and sentence recall tasks (r (96) = .63, p < .001). Therefore, these four tasks were consolidated into two composite distal language measures reflecting words (composite oral vocabulary score consisting of the expressive and receptive vocabulary tasks) and sentences (composite oral 'sentence' score consisting of the sentence generations and sentence recall tasks). Composite measures were used in all the analyses.

Working memory. The listening span task from Henry (2001) was used. Children hear implausible/sensible sentences and asked to judge whether they are true or false while also trying to remember the final word of the sentence. The task increases in difficulty and the

final score obtained is the longest list length where two trials were passed, plus a half-point credit if one trial at the next list length was correctly recalled (minimum score 0, maximum score 4). Reported reliability is r = .76 and validity r = .74.

Single Word Reading. The Single Word Reading Test (SWRT; Foster, 2007), in which children read a list of words, provided a measure of untimed word reading accuracy.

Reading Comprehension. The YARC (Snowling et al., 2009) provided a measure of reading comprehension. Children read the passages determined by their word reading level and answer a series of open-ended literal and inferential questions about the text. Scores are transformed into ability scores to allow comparison across the passages.

Non-verbal ability. The matrices task (BAS II; Elliott et al., 1997) was used. Participants select a picture that completes a pattern from a set of six. Test-retest reliability is r = .64 and validity is r = .47.

2.3 Procedure

Children were tested individually in a quiet room at school over three sessions. Standardised measures were administered first. Children performed the CBM writing task on a digitizing tablet (100 Hz, Intuos 4; Wacom, Vancouver, Washington) and recorded using Eye & Pen software. Participants write on lined paper (taped onto the tablet) with an inking pen. The tablet surface records the xy coordinates of the pen's position to a Windows based computer.

Writing product coding for the CBM Writing task

Productivity was measured by counting the number of words, irrespective of spelling accuracy, produced in the five-minute period, excluding the prompt, numerals and crossed-out words. Accuracy was assessed by counting correct word sequences (CWS) and subtracting the number of incorrect word sequences (ICWS) (Romig, Therrien, & Lloyd, 2017). CWS were defined as any two adjacent words that were acceptable within the context

of the sample. End punctuation and beginning capitalization were included in the scoring of CWS. Examples of scoring for the word sequences measure can be found in Dockrell et al. (2015). To distinguish transcription errors and text generation skills, spelling errors were not penalized. Text quality was assessed using a modification of the WIAT (Weschler, 2005) paragraph scoring (0 to 6) ignoring spelling grammar and punctuation, where 0 was an unintelligible text or too few words to judge the content of the text and 6 was a well written text that presented a clear, organized and developed description of the topic.

The scripts were typed and guidelines for scoring created. Reliability checks were performed on all the anonymized written texts by two raters unfamiliar with the participants but trained in assessing written texts. Interrater reliability evaluated through intra-class correlations were high: total number of words 1.0, total number of words spelled correctly 1.0, CWS .98, and text quality .86.

Writing process analysis from the Eye and Pen software temporal data

The handwritten scripts temporal data were analyzed into writing bursts. Following common convention in this research area (Alves & Limpo, 2015; Chenoweth & Hayes, 2001). A burst was defined as a period of writing activity between two 2 second or longer consecutive pauses in which at least one word was written and burst length as the number of words written in the given burst. A mean burst length score was generated for each child from the burst length data produced during the CBM task.

3. Results

The results are presented in three sections. First, we describe the performance of the group of struggling writers identified using the standardised WOLD writing measure on our writing product measures and the measure of writing bursts. Differences between participants in proximal and distal measures are also described to assess the pervasiveness of their problems. As described in the methods two composite oral language measures were created

reflecting oral vocabulary performance and oral 'sentence' performance. We transformed all tests results to a population Z score (Mean of 0, SD of 1) to allow comparisons across the measures based on different standard scales. Raw scores were used for all measures where participants completed the same items and ability scores where tests differentiated by age (spelling and reading comprehension).

In section 2 ROC analysis is used to determine the most sensitive and specific discriminator of the SWs (Keller-Margulis, et al., 2016) from the writing product and process measures described in section 1 and logistic regression is used to identify the statistically significant discriminator variables. Section 3 establishes which proximal and distal measures account for variation in performance on the discriminator variable(s) identified in section 2. Path analyses were then used to explore the direct and indirect relationships between proximal and distal factors and the writing product.

Section 3.1: Writing performance and cognitive underpinnings

Table 1 provides means (SDs) for the population Z scores for the two groups of writers (SW and TW) for all the writing measures and for proximal and distal factors identified by the literature. These results confirm the grouping of SWs and TWs identified by the WOLD writing assessment. As the table shows, despite significant variability for both groups, on all measures the mean Z scores for SWs was negative. Raw scores are presented in Appendix 1.

SWs performed significantly more poorly than TWs on all the measures of written production examined. A MANCOVA, controlling for non-verbal ability, compared the two groups' performance on all the writing measures. There was a statistically significant difference by group, $F_{(4, 90)} = 8.50$, $p < .001$; Wilk's $\Lambda = 0.76$, $\eta_p^2 = .27$, and non-verbal ability, $F_{(4, 90)} = 4.99$, $p = .001$; Wilk's $\Lambda = 0.81$, $\eta_p^2 =. 18$. Subsequent univariate testing using ANCOVA controlling for non-verbal ability indicated that SWs were significantly

poorer on all the product measures from the CBM task (productivity, $F (2, 95) = 19.41$, $p <$ .001, $\eta_p^2 = .17$; text accuracy, $F (2, 95) = 21.97$, $p < .001$, $\eta_p^2 = .19$; text quality, $F (2, 95) = 24.16$, $p < .001$, $\eta_p^2 = .21$) and in the written sentence generation task, $F (2, 95) = 26.26$, $p < .001$, $\eta_p^2 = .22$. In all cases, effect sizes were large. The two groups also differed significantly on the writing process measures and, as predicted, SWs produced shorter writing bursts, $F (2, 95) = 14.29$, $p < .001$, $\eta_p^2 = .13$, although the effect size was only medium (see Cohen, 1988).

As Table 1 shows, SWs were also weaker in the hypothesised proximal and distal measures that underpin written text production. A MANCOVA, controlling for non-verbal ability, compared the two groups' performance on these measures. There was a statistically significant difference by group, $F (5, 89) = 10.47$, $p < .001$; Wilk's $\Lambda = 0.63$, $\eta_p^2 = .37$, and non- verbal ability, $F (5, 89) = 3.21$, $p = .01$; Wilk's $\Lambda = 0.85$, $\eta_p^2 =. 15$. Subsequent univariate testing using ANCOVA controlling for non-verbal ability indicated that SWs were significantly poorer on both proximal measures of writing: spelling, $F (1, 93) = 38.87$, $p < .001$, $\eta_p^2 = .30$; handwriting fluency, $F (1, 93) = 19.03$, $p < .001$, $\eta_p^2 = .17$. SWs also performed more poorly on distal measures of writing: oral vocabulary, $F (1, 93) = 27.88$, $p < .001$, $\eta_p^2 = .23$; oral 'sentence' scores, $F (1, 93) = 28.89$, $p < .001$, $\eta_p^2 = .24$; reading decoding, $F (1, 93) = 45.04$, $p < .001$, $\eta_p^2 = .33$; reading comprehension, $F (1, 93) = 19.10$, $p < .001$, $\eta_p^2 = .17$; and listening span, $F (1, 93) = 11.69$, $p < .001$, $\eta_p^2 = .11$. Nonverbal ability was only significant in the models for oral vocabulary, $F (1, 93) = 13.62$, $p < .001$, $\eta_p^2 = .13$, and the oral 'sentence' scores, $F (1, 93) = 7.73$, $p = .007$, $\eta_p^2 = .08$. All other $ps > .3$.

In sum, the identification of the group of SWs was confirmed in separate measures of writing at word, sentence and text levels and also in burst fluency. In addition, SWs experienced difficulties with both proximal and distal competencies that support the production of written text. Effect sizes for spelling, oral language and reading decoding were

large. By contrast, effect sizes for handwriting fluency, reading comprehension and listening span were smaller.

Section 3.2: Which writing measure best discriminates struggling writers?

We predicted that SWs and TWs would best be distinguished by the quality of their written product and the writing process measure (writing bursts). To test this prediction, a ROC analysis was computed (Keller-Margulis et al., 2016) to establish which writing measures (see Table 1) most accurately identified the SWs. The area under the curve and 95 per cent confidence intervals are presented in Table 2. As Table 2 shows, all measures were significant and, the 95% confidence intervals do not include the null hypotheses of .5. Benchmarks for gauging AUCs (Swets, Dawes & Monahan, 2000) indicate that word productivity, grammatical accuracy, text quality and writing bursts in words demonstrated fair ($\geq$ .70) sensitivity and specificity in identifying the SWs at this age. Only the written sentence generation task demonstrated the good sensitivity and specificity required to be defined as a reliable discriminator ($\geq$ .80). Therefore, on average, a child scoring poorly on the written sentence generation is more likely than 83% of the controls to be identified as a SW.

A logistic regression analysis further tested whether written sentence generation was significantly more likely to predict group membership than the alternate writing measures. A test of the full model against a constant only model was statistically significant indicating that the predictors, as a set, reliably distinguished SWs and TWs (X^2 (5, N = 96) = 36.44, p < .001). Nagelkerke's R^2 of .43 indicated a moderate relationship between prediction and grouping. Prediction success overall was 78% (80% for TWs and 74% for SWs). The Wald criterion (7.92) demonstrated that, again, only the written sentence generation task made a significant contribution to prediction (p = .005). The *Exp(B)* value (.83) indicated that as the sentence generation scores increased, the odds of being classified as a SW reduced by 17 per

cent. For pupils in middle elementary school the written sentence generation task provided the best measure to discriminate SWs in our sample compared to the other writing measures.

Section 3.3: Proximal and distal factors of writing

We first explored the zero order correlations between writing measures and the proximal and distal measures collected (see Table 3). Bonferroni's adjustment (p =.001 for significance) was used to control for multiple correlations. As expected there were large and significant correlations between the writing measures, with writing productivity and accuracy demonstrating the largest association. Writing product measures showed higher correlations with both proximal and distal factors than the writing burst measure. Only the writing burst measure was significantly associated with non-verbal ability.

As described in section 2, the written sentence generation task, and not the other measures of writing assessed, was the only significant writing measure that statistically distinguished the groups of SWs and their TWs using logistic regression. Thus, to further elucidate the differences between the SWs and the TWs the focus of our further analyses was on the proximal and distal predictors of the written sentence generation task alone.

Multiple linear regression analysis was used to develop a model for predicting written sentence generation performance from performance on the proximal and distal measures assessed. It is evident in Table 3 that, apart from non-verbal ability, each of the predictor variables had a significant zero-order correlation with written sentence generation performance. Given the collinearity between single word reading and single spelling (r = .88) only single word spelling was included in the regression. A stepwise multiple linear regression was calculated to establish the impact of the proximal and distal measures on written sentence generation. Age was entered in the first step, followed by the proximal predictors of writing (handwriting fluency and spelling) and in the next step the distal measures of listening span, oral language (composite oral vocabulary scores and oral

'sentence' scores) and reading comprehension. Each additional step in the regression resulted in a statistically significant change. The final model included as significant predictors, handwriting fluency, oral 'sentence' scores and listening span, $F(6, 95) = 21.65, p < .001, R^2 = .63, R^2_{Adjusted} = .60$. Table 4 presents the results of the final model.

We further tested our results using path analysis to establish proximal (direct) and distal (indirect) contributions to the written sentence generation task. Path analysis is theoretically useful because, unlike other techniques, it forces us to specify relationships among all of the independent variables. This results in a model showing mechanisms through which independent variables produce direct and indirect contributions on a dependent variable. The overall fit of the final models were assessed by x^2 and CFI. Following Hu and Bentler (1999), who recommend joint criteria to retain a model: models were only considered a good fit if the x^2 was not significant, RMR was close to zero and CFI > .96, CFI being a more sensitive fit index with small sample sizes.

As described in the introduction written sentence generation taps the demands of text generation and transcription. We reasoned from the current literature examining writing that both oral 'sentence' skills and WM (measured by listening span) were distal factors, i.e. exogenous variables, and spelling and handwriting fluency proximal, that is endogenous variables in the model. We thus first tested a model where the distal factors, oral 'sentence' generation and listening span would work indirectly through spelling on written sentence generation, and both spelling and handwriting fluency (proximal factors) would work directly on written sentence generation. This model tested the hypothesis that these distal factors of writing performance worked through the hypothesised building blocks of writing, that is, through the proximal factors (spelling and handwriting fluency). This model had a poor fit ($X^2(3) = 18.03, p < .001$, RMR .25, CFI = .71). We also considered a variant of this model

where spelling and handwriting fluency were considered as covariates. This model also had a poor fit ($X^2(3) = 12.08$, $p < .001$, RMR .05, CFI = .92).

We conducted a final test where, based on Savage et al. (2017), we explored the possibility that written sentence generation would be predicted by oral 'sentence' generation both directly and indirectly through spelling. We further reasoned that by this point in development spelling accuracy would predict handwriting fluency (Abbott, Berninger, Fayol, 2010). This model was significant. The goodness of fit measures indicated a good fit: $X^2(3) = 7.22$, $p = .07$, RMR = .035, CF1 = .98. The final model explained 59 per cent of the variance in the written sentence generation task. Standardised betas are presented to provide comparable data for the contribution of the variables. As figure 1 shows, oral 'sentence' skills were statistically significantly related directly to written sentence levels skills ($\beta = .35$) and spelling ($\beta = .63$). The role of listening span on written sentence generation was not significant, but correlated highly with oral 'sentence' skills. In addition, both spelling ($\beta = .29$) and handwriting fluency ($\beta = .34$) directly supported written sentence generation. Oral 'sentence' skills were also statistically significantly indirectly related to written sentence level skills through spelling ($\beta = .31$) and handwriting fluency through ($\beta = .34$). Finally spelling was indirectly related to written sentence level skills through handwriting fluency ($\beta = .21$). In sum, the model supported and extended the regression analyses by demonstrating that oral 'sentence' performance, as a distal factor, both directly and indirectly contributed to written sentence generation, and spelling and handwriting fluency, as proximal factors significantly contributed to written sentence generation.

4. DISCUSSION

We conducted a study to examine which writing measure best discriminated SWs and their age matched peers in middle elementary school and to examine the proximal and distal factors that underpinned performance on the writing measure that best disriminated SWs. SWs were initially identified on a standardised norm referenced writing measure and their performance further evaluated on independent writing assessments examining word productivity, written sentence generation, grammatical accuracy, text level quality as well as writing bursts. Given the research literature we predicted that the quality of the children's written text and their burst length, a measure of the fluency of compositional process, would be the measures that were the most accurate discriminators of SWs. Contrary to our predictions it was only a written sentence generation task that provided a high level of diagnostic accuracy.

Consistent with extant frameworks of the development of writing, SWs were weaker on both their proximal skills (poor spelling and handwriting fluency) and on distal skills (oral language, reading and listening span), which underpin the production of written text. We reasoned that both proximal and distal factors would contribute to performance on the best discriminator variable, written sentence generation. We predicted, given the research literature, that spelling would serve as the proximal variable and oral 'sentence' generation the distal measure explaining variance in the task. As predicted, oral 'sentence' generation was a significant distal factor, but handwriting fluency was the proximal predictor. Contrary to our expectations that distal factors would impact indirectly through proximal factors the path analysis demonstrated, oral 'sentence' generation worked both indirectly through spelling and directly to written sentence generation.

4.1 Sentence level skills and writing

The current study extends our understanding of the importance and the skills that underpin written sentence production. Using a written sentence generation task, thereby

reducing the potential impact of other cognitive factors, allowed us to isolate the diagnostic significance of sentence level performance in children who struggle with writing. In middle elementary school as children develop automaticity in their transcription skills then sentence levels skills contribute uniquely to their ability to translate ideas into text (Berninger et al., 2011). Creating sentences becomes a problem space in its own right for these children, drawing on grammatical rules and punctuation (Shanahan, 2006), as well as the ability to co-ordinate ideas. The data indicate that SWs have significant difficulties in textual production at the sentence level; where the ability to formulate accurate sentences appears to be a necessary step in developing wider writing competencies (Berninger, et al., 1992). Although current descriptions of writing often refer to word, sentence and text level skills much published research on children learning to write focuses on text level performance, followed by a substantial body of work focusing on transcription. Rarely is there an attempt to isolate sentence level competency. Our results point to the importance of sentence level performance and the need to capture factors such as sentence generation that underpin writing performance. The results extend previous studies (Arfé & Pizzoccaro, 2016; Arfé et al., 2016; Berninger et al., 2011) that illustrate the possibility that competence at the sentence level is an important developmental hurdle that children need to achieve in order to create quality text level writing. Further work needs to provide evidence to develop our understanding of the role of sentence level skills.

4.2 Proximal and distal factors supporting sentence generation

Following the research on reading, we drew a conceptual distinction between proximal and distal factors and this was tested empirically in our path models. Our final model identified the importance of proximal and distal factors and their direct and indirect contribution to written sentence generation. In the path model spelling, as a proximal factor, had both a direct and indirect contribution to written sentence generation. By contrast

handwriting fluency only contributed directly. These results confirm the continued importance of transcription skills as proximal factors in writing production in middle elementary school. Contrary to the prediction of distal factors working indirectly oral 'sentence' generation, our theorised distal measure, contributed both directly to written sentence generation. Listening span, while correlated with oral 'sentence' generation did not contribute either directly or indirectly to the path model. Two possible explanations of this non-significant result are considered. The Listening span task is similar to sentence processing task and to capture independent distal contributions of WM it may be necessary to use different measures. Conceptually, it may be that reduced demands in creating links throughout a text reduces the WM demands for the writer. Further research is needed to elucidate potential explanations.

We had anticipated that by this point in development handwriting fluency would be sufficiently automatic that it would reduce in importance to allow us to illustrate how other skills support the writing process. This was only partially confirmed and while oral 'sentence' generation and listening span explained variance in written sentence generation in the regression analyses, fluency in handwriting continued to have a significant impact at the sentence level. This supports research on handwriting fluency (Graham, et al., 1997; Alves & Limpo, 2015) but it may be that the sentence generation task accentuated this result. The task provided the key word pairs written on the page, reducing spelling demands. Additionally, the task was time limited with children instructed to write as many sentences as possible, likely accentuating the impact of speed of letter writing, which was the measure of handwriting fluency.

Oral 'sentence' generation was also a predictor of written sentence generation. Similar results from other researchers, which speak to the importance of oral 'sentence' level skills over time, have been found (Savage et al., 2007) and oral 'sentence' skills predicted written

grammatical skills in children with language impairment and Autism Spectrum Disorder (Dockrell, Ricketts, Charman & Lindsay, 2014). The impact of oral language at the sentence level is of theoretical import. Previous research has demonstrated how vocabulary skills and narrative skills concurrently and over time support text level generation (Dockrell et al., 2009). Here we show that oral 'sentence' generation skills are strongly associated with written sentence generation skills and act independently of transcription skills. Further variance was explained by listening span, our measure of WM, supporting arguments highlighting the distal role of WM in sentence generation tasks (Olive & Cislaru, 2015).

Studies examining the ways in which oral language underpin written text generation need to consider the links at word, sentence and text level (Ahmed et al., 2014). The current path analysis indicated that oral 'sentence' level skills predicted written sentence level performance both directly and indirectly through spelling. This dual contribution of our distal measures highlights the importance of further examining the ways in which the skills, which underpin writing, interact and predict different elements of writing.

4.3 Practical implications

This study has illustrated that the sentence level may be a reliable and valid way to make an assessment of writing ability in middle elementary school. The greater awareness of the development of written text production and the cognitive processes which support written text production has led to an increased emphasis on the adaptations teachers can make to support SWs (Graham, Harris, Bartlett, Popadopoulou, & Santoro, 2016). While the importance of sentence level skills have been emphasised in this study, further research is required to determine how to best translate this knowledge into practice. There is evidence that students benefit from sentence level work to improve overall writing quality (Saddler, Behforooz, & Asaro, 2008), further highlighting the importance of sentence level skills. A series of small-scale interventions to improve sentence construction in middle school students

has also had some success with sentence construction (Datchuk, 2016) and provides the basis to develop future work.

4.4 Limitations

This was a correlational study and so no direct causal relations can be established. The modelling was exploratory and would need to be tested in further studies. We did not measure the wide variety of distal skills that have been implicated in the development of writing such as executive skills or motivational factors. Nor did we consider the learning environment (MacArthur & Graham, 2016). There was also no analysis of any planning or revision aspects involved in writing (though there is little direct evidence planning or revision is done in any detail at this age, Singer & Bashir, 2004). The use of single CBM tasks to assess writing has been challenged (see Gansle, VanDerHeyden, Noell, Resetar, & Williams, 2006 for a review) and concerns about brevity could be levelled at our written sentence generation task. Nonetheless, we used a longer independent standardised writing assessment to classify the groups and thus this is one of the few studies to capture performance deficits in SWs using multiple measures of writing within samples, as recommended by recent studies (Bouwer, Beguin, Sanders & Van den Bergh, 2014; Rietdijk, Janssen, van Weijen, van der Bergh & Rijlaarsdam, 2017).

4.3 Summary

The inability to produce sustained, accurate, and competent writing remains a pervasive weakness for many students; teachers need to be able to profile pupils developing writing skills so appropriate action can be taken. This study develops the emerging evidence that SWs have difficulties in textual production at the sentence level (Garcia et al., 2017) and that a test of written sentence generation discriminated between SWs and TWs in middle elementary school. Using a conceptual model which captured proximal and distal skills which support the writing process we demonstrated both direct and indirect contributions on

written sentence level production. The proximal measures of handwriting fluency and the

distal measures of oral language sentence generation and listening span predicted

performance on the written sentence generation task in the regression analyses. The path

analysis further developed the framework by identifying the contribution of oral 'sentence'

generation skills to written sentence generation skills directly. The results underscore the

continuing importance of transcription in middle elementary school but additionally the role

of oral language in the development of writing at the sentence level.

REFERENCES

Abbott, R. D., & Berninger, V. W. (1993). Structural Equation Modeling of Relationships Among Developmental Skills and Writing Skills in Primary-Grade and Intermediate - Grade Writers. *Journal of Educational Psychology, 85*(3). doi:10.1037/0022-0663.85.3.478

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal Relationships of Levels of Language in Writing and Between Writing and Reading in Grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281-298. doi:10.1037/a0019318

Ahmed, Y., Wagner, R. K., & Lopez, D. (2014). Developmental Relations Between Reading and Writing at the Word, Sentence, and Text Levels: A Latent Change Score Analysis. *Journal of Educational Psychology, 106*(2), 419-434. doi:10.1037/a0035692

Alves, R. A., & Limpo, T. (2015). Progress in Written Language Bursts, Pauses, Transcription, and Written Composition Across Schooling. *Scientific Studies of Reading, 19*(5), 374-391. doi:10.1080/10888438.2015.1059838

Arfé, B., Dockrell, J. E., & De Bernardi, B. (2016). The effect of language specific factors on early written composition: the role of spelling, oral language and text generation skills in a shallow orthography. *Reading and Writing, 29*(3), 501-527. doi:10.1007/s11145-015-9617-5

Arfé B., & Pizzocaro, E. (2016). Sentence generation in children with and without problems of written expression. In J. Perera, M. Aparici, E. Rosado, & N. Salas (Eds.), *Written and Spoken Development across the Lifespan*. New York: Springer.

Barnett, A., Henderson, S. E., Scheib, B., & Schulz, J. (2007). Detailed Assessment of Speed of Handwriting (DASH). In. London: Harcourt Assessment.

Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-Level Developmental Skills in Beginning Writing. *Reading and Writing, 4*(3), 257-280. doi:10.1007/bf01027151

Berninger, V. W., & Swanson, H. L. (1994). Modifying Hayes and Flowers' model of skilled writing to explain developing writing. In E. C. Butterfield (Ed.), *Advances in cognition and educational practice. Children's writing: Toward a process theory of the development of skilled writing* (Vol. 2, pp. 1-30). Greenwich: JAI Press.

Berninger, V. W., & Winn, W. D. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In

C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research*. New York: Guildford.

Berninger, V. W., Nagy, W., & Beers, S. (2011). Child writers' construction and reconstruction of single sentences and construction of multi-sentence texts: contributions of syntax and transcription to translation. *Reading and Writing, 24*(2), 151-182. doi:10.1007/s11145-010-9262-y

Bouwer, R., Béguin, A., Sanders, T., & Van den Bergh, H. (2014). Effect of genre on the generalizability of writing scores. *Language Testing, 32*(1), 83-100. doi: 10.1177/0265532214 542994

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing - Generating text in L1 and L2. *Written Communication, 18*(1), 80-98. doi:10.1177/0741088301018001004

Cohen J. (1988). Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge Academic

Connelly, V., Dockrell, J. E., & Barnett, A. (2012). Children Challenged by Writing Due to Language and Motor Difficulties. In V. Berninger (Ed.), *Cognitive Psychology of Writing Handbook: Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*. New York; NY: Psychology Press.

Costa, J., Green, M., Sideris, J. & Hooper, S.R. (2017). First-Grade Cognitive Predictors or Writing Disabilities in Grades 2 through 4 Elementary School Students. *Journal of Learning Disabilities* **First Published** July 18, 2017 https://doi.org/10.1177/0022219417721182

Cragg, L., & Nation, K. (2006). Exploring written narrative in children with poor reading comprehension. *Educational Psychology, 26*, 17.

Datchuk, S. (2016). Writing simple sentences and descriptive para- graphs: Effects of an intervention on adolescents with writing difficulties. *Journal of Behavioral Education*, *25*, 166–18

Dockrell, J. E., Connelly, V., Walter, K., & Critten, S. (2015). Assessing children's writing products: the role of curriculum based measures. *British Educational Research Journal, 41*(4), 575-595. doi:10.1002/berj.3162

Dockrell, J. E., Lindsay, G., & Connelly, V. (2009). The Impact of Specific Language Impairment on Adolescents' Written Text. *Exceptional Children, 75*(4), 427-446.

Dockrell, J. E., Ricketts, J., Charman, T., & Lindsay, G. (2014). Exploring writing products in students with language impairments and autism spectrum disorders. *Learning and Instruction, 32*, 81-90. doi:10.1016/j.learninstruc.2014.01.008

Dunn, L. M., & Dunn, D. M. (2009). British Picture Vocabulary Scale III (BPVS-III). In. UK: GL Assessments.

Elliott, C. D., Smith, P., & McCulloch, K. (1997). The British Ability Scales II. In. Windsor, UK: NFER-Nelson.

Foster, H. (2007). Single word reading test 6–16. London: GL Assessment.

Gansle, K. A., VanDerHeyden, A. M., Noell, G. H., Resetar, J. L., & Williams, K. L. (2006). The technical adequacy of curriculum-based and rating-based measures of written expression for elementary school students. *School Psychology Review, 35*(3), 435-450.

Garcia, E., Crespo, P., & Bermudez, I. (2017). Writing an Independently Composed Sentence by Spanish-Speaking Children With and Without Poor Transcription Skills: A Writing-Level Match Design. *Journal of Learning Disabilities, 50*(5), 511-521. doi:10.1177/0022219416633862

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students. A new methodological

approach. *Journal of Educational Psychology, 89*(1). doi:10.1037/0022-0663.89.1.170

Graham, S., & Harris, K. R. (2005). Improving the writing performance of young struggling writers: Theoretical and programmatic research from the center on accelerating student learning. *Journal of Special Education, 39*(1). doi:10.1177/00224669050390010301

Graham, S., Harris, K. R., Bartlett, B. J., Popadopoulou, E., & Santoro, J. (2016). Acceptability of Adaptations for Struggling Writers: A National Survey With Primary-Grade Teachers. *Learning Disability Quarterly, 39*(1), 5-16. doi:10.1177/0731948714554038

Graham, S., Harris, K. R., Fink-Chorzempa, B., & MacArthur, C. (2003). Primary grade teachers' instructional adaptations for struggling writers: A national survey. *Journal of Educational Psychology, 95*(2), 279-292. doi:10.1037/0022-0663.95.2.279

Graham, S., Hebert, M., & Harris, K. (2011). Throw 'em Out or Make 'em Better? State and District High-Stakes Writing Assessments. *Focus on Exceptional Children, 44*(1).

Hayes, J. R. (2012). Modeling and Remodeling Writing. *Written Communication, 29*(3), 369-388. doi:10.1177/0741088312451260

Henry, L. (2001). How does the severity of a learning disability affect working memory performance? *Memory*, 9, 233-247

He, Q. P., Anwyll, S., Glanville, M., & Deavall, A. (2013). An investigation of the reliability of marking of the Key Stage 2 National Curriculum English writing tests in England. *Educational Research, 55*(4), 393-410. doi:10.1080/00131881.2013.844942

Keller-Margulis, M., Payan, A., Jaspers, K. E., & Brewton, C. (2016). Validity and Diagnostic Accuracy of Written Expression Curriculum-Based Measurement for Students With Diverse Language Backgrounds. *Reading & Writing Quarterly, 32*(2), 174-198. doi:10.1080/10573569.2014.964352

Kent, S. C., & Wanzek, J. (2016). The Relationship Between Component Skills and Writing Quality and Production Across Developmental Levels: A Meta-Analysis of the Last 25 Years. *Review of Educational Research, 86*(2), 570-601. doi:10.3102/0034654315619491

Kim, Y.-S., Al Otaiba, S., Folsom, J. S., Greulich, L., & Puranik, C. (2014). Evaluating the Dimensionality of First-Grade Written Composition. *Journal of Speech Language and Hearing Research, 57*(1), 199-211. doi:10.1044/1092-4388(2013/12-0152)

Kim, Y. S., Al Otaiba, S., & Wanzek, J. (2015). Kindergarten predictors of third grade writing. *Learning and Individual Differences, 37*, 27-37. doi:10.1016/j.lindif.2014.11.009

Kim, Y. K., Gaitlin, B., Al Otaiba, S., & Wanzek, J. (2017). Theorization and an Empirical Investigation of the Component-Based and Developmental Text Writing Fluency Construct. *Journal of Learning Disabilities*. doi:10.1177/0022219417712016

Kim, Y. S. G., & Schatschneider, C. (2017). Expanding the Developmental Models of Writing: A Direct and Indirect Effects Model of Developmental Writing (DIEW). *Journal of Educational Psychology, 109*(1), 35-50. doi:10.1037/edu0000129

MacArthur, C., & Graham, S. (2016). Writing Research from a Cognitive Perspective. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 24-40). New York: Guildford Publications.

Mackie, C. J., Dockrell, J.E., & Lindsay, G. (2013). An evaluation of the written texts of children with SLI: the contributions of oral language, reading and phonological short-term memory. *Reading and Writing, 26*(6), 865-888. doi:10.1007/s11145-012-9396-1

McCutchen, D. (1986). Domain Knowledge and Linguistic Knowledge in the Development of Writing Ability *Journal of Memory and Language, 25*(4), 431-444.

McCutchen, D. (2012). Phonological, orthographic, and morphological word-level skills supporting multiple levels of the writing process. In V. W. Berninger (Ed.) *Past, present and future contributions of cognitive writing research to cognitive psychology* (pp. 197-216). New York, NY: Psychology Press.

McCutchen, D., Covill, A., Hoyne, S. H., & Mildes, K. (1994). Individual Differences in Writing – Implications for Translating Fluency. *Journal of Educational Psychology, 86*(2). doi:10.1037//0022-0663.86.2.256

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing - A literature review. *Journal of Special Education, 41*(2). doi:10.1177/00224669070410020301

Olinghouse, N. G., & Graham, S. (2009). The Relationship Between the Discourse Knowledge and the Writing Performance of Elementary-Grade Students. *Journal of Educational Psychology, 101*(1), 37-50. doi:10.1037/a0013248

Olinghouse, N. G., & Leaird, J. T. (2009). The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and Writing, 22*(5), 545-565. doi:10.1007/s11145-008-9124-z

Olinghouse, N. G., & Santangelo, T. (2010). Assessing the Writing of Struggling Learners. *Focus on Exceptional Children, 43*(4), 1-27.

Olive, T., & Cislaru, G. (2015). Linguistic forms at the product process interface: Analysing the linguistic content of bursts of production. In G. Cislaru (Ed.), *Writing(s) at the crossroads: The product process interface* (pp. 99-124). Amsterdam: John Benjamins Publishing Company.

Olive, T., Favart, M., Beauvais, C., & Beauvais, L. (2009). Children's cognitive effort and fluency in writing: Effects of genre and of handwriting automatisation. *Learning and Instruction, 19*(4), 299-308. doi:10.1016/j.learninstruc.2008.05.005

Rietdijk, S., Janssen, T., Van Weijen, D., Van den Bergh, H., & Rijlaarsdam, G. (2017). Improving writing in primary schools through a comprehensive writing program. *Journal of Writing Research, 9*(2)*,* 173-225. doi: 10.17239/jowr-2017.09.02.04

Romig, J. E., Therrien, W. J., & Lloyd, J. W. (2017). Meta-Analysis of Criterion Validity for Curriculum-Based Measurement in Written Language. *Journal of Special Education, 51*(2), 72-82. doi:10.1177/0022466916670637

Rust, J. (1996). *The manual of the Wechsler objective language dimensions (WOLD) UK Edition*. London: The Psychological Corporation.

Saddler, B., Behforooz, B., & Asaro, K. (2008). The Effects of Sentence-Combining Instruction on the Writing of Fourth-Grade Students With Writing Difficulties. *Journal of Special Education, 42*(2), 79-90. doi:10.1177/0022466907310371

Savage, R., Kozakewich, M., Genesee, F., Erdos, C., & Haigh, C. (2017). Predicting writing development in dual language instructional contexts: exploring cross-linguistic relationships. *Developmental Science, 20*(1). doi:10.1111/desc.12406

Semel, E., Wiig, E., & Secord, W. (2006). Clinical Evaluation of Language Fundamentals - Fourth Edition (CELF - 4 UK). In: Pearson Assessments.

Shanahan, T. (2006). Relations among oral language, reading and writing development. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 171-183). New York: Guildford Press.

Singer, B. D., & Bashir, A. S. (2004). Developmental variations in writing composition skills. In Stone, C. A., Silliman, E. R., Ehren, B. J., & Apel, K.. (Eds.). *Handbook of language and literacy: Development and disorders*. (pp. 559-582), New York, Guilford Press.

Snowling, M.J., Stothard, S.E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., & Nation, K. (2009). York assessment of reading for comprehension: Passage reading. London: GL Assessment

Swanson, H. L. (1996). Individual and age-related differences in children's working memory. *Memory & Cognition, 24*(1), 70-82.

Swets, J.A., Dawes, R.M. & Monahan, J. (2000). Psychological Science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1* 1-26.

Wechsler, D. (2005). Wechsler Individual Achievement Test (WIAT-II UK). In. London: Pearson; Harcourt Assessments.

Williams, G. J., Larkin, R. F., & Blaggan, S. (2013). Written language skills in children with specific language impairment. *International Journal of Language & Communication Disorders, 48*(2), 160-171. doi:10.1111/1460-6984.12010

Zoccolotti, P., De Luca, M., Marinelli, C. V., & Spinelli, D. (2014). Modeling individual differences in text reading fluency: a different pattern of predictors for typically developing and dyslexic readers. *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.01374
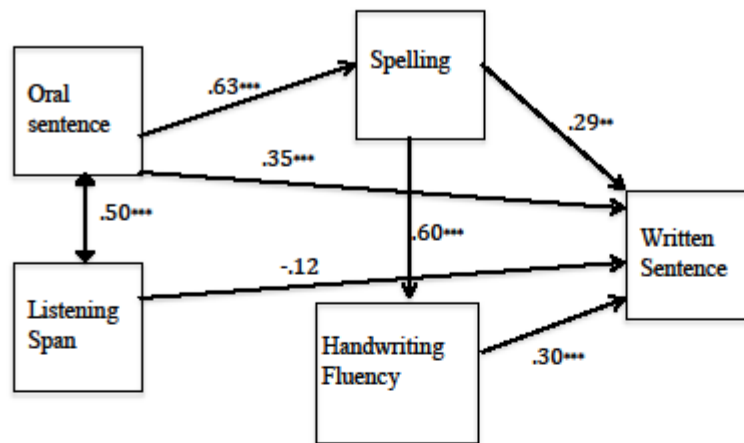
Figure 1. Standardised regression coefficients and significance levels for proximal and distal factors and written sentence production

** p < .01, *** p <.001

Table 1 Mean Z scores (SDs) for typical writers and struggling writers on writing outcomes, proximal and distal factors

	Typical writers	Struggling writers
	Mean (SD)	Mean (SD)
Writing measures		
Productivity (numbers of words written)	.30 (.92)	-.44 (.95)
Accuracy (Correct Word Sequences-Incorrect word sequences)	.35 (.99)	-.51 (.78)
Quality	.37 (.94)	-.54 (.83)
Written sentence generation	.41 (.95)	-.60 (.74)
Writing process (bursts)	.27 (1.02)	-.38 (.83)
Proximal factors		
Spelling	.46 (.88)	-.67 (.75)
Handwriting fluency	.33 (1.04)	-.48 (.72)
Distal factors		
Listening Span (WM)	.31 (.95)	-.45 (.91)
Oral vocabulary	.43 (.88)	-.63 (.82)
Oral 'sentence'	.43 (.93)	-.63 (.72)
Single word reading	.49 (.72)	-.71 (.93)
Reading comprehension	.36 (1.0)	-.53 (.73

Table 2 ROC analysis for writing product and process measures

Writing Measure		Area under curve (AUC)	Sig	95% Confidence Interval Lower Bound	Upper Bound
Product measures	Productivity	0.72	.001	0.61	0.82
	Accuracy	0.74	.001	0.64	0.84
	Quality	0.76	.001	0.66	0.86
	Written Sentence Generation	0.83	.001	0.74	0.91
Process measure	Average writing burst length in words	0.74	.001	0.64	0.84

Table 3 Zero order correlations between writing measures, proximal and distal factors and non-verbal ability.

	Written Language measures					Proximal factors			Distal factors			
	Written Sentence generation.	Productivity	Accuracy	Quality	Writing bursts	Spelling	Handwriting fluency	Single word reading	Reading comprehension	Oral vocabulary	Oral 'Sentence'	Listening Span (WM)
Productivity	.57**											
Accuracy	.64**	.92**										
Quality	.56**	.73**	.75**									
Writing bursts	.46**	.80**	.73**	.50**								
Spelling	.66**	.61**	.72**	.65**	.47**							
Handwriting fluency	.66**	.62**	.63**	.53**	.55**	.60**						
Single word reading	.61**	.51**	.64**	.57**	.35**	.88**	.51**					
Reading comprehension	.58**	.49**	.59**	.55**	.30**	.75**	.46**	.76**				
Oral vocabulary	.52**	.26	.35**	.44**	.13	.57**	.35**	.59**	.52**			
Oral 'sentence'	.63**	.43**	.55**	.54**	.34*	.63**	.48**	.61**	.55**	.73**		
Listening Span (WM)	.34*	.14	.20	.31	.08	.37**	.13	.39**	.35**	.45**	.50**	
Non Verbal Ability	.24	-.13	.01	.05	.47**	.17	.08	.20	.21	.42**	.36**	.23

N = 96, Bonferroni's adjustment: * p =. 001, ** p< .0005

Table 4   Regressions examining predictors of written sentence generation

	B	Std error	Beta	t	sig
Chronological age	-0.114	0.06	-0.133	-1.918	*ns*
Spelling	0.113	0.042	0.259	2.681	*ns*
Handwriting fluency	0.211	0.05	0.354	4.176	.0005
Listening span	-1.686	0.847	-0.157	-1.99	.04
Oral vocabulary	0.046	0.045	0.101	1.033	*ns*
Oral 'sentence'	0.15	0.05	0.336	3.018	.004
Reading comprehension	0.13	0.09	0.144	1.45	*ns*

Appendix 1 Mean raw scores (SDs) for typical writers and struggling writers on writing outcomes, proximal and distal factors

		Typical writers	Struggling writers
		Mean (*SD*)	Mean (*SD*)
Writing			
Product measures	Productivity (number of words written)	68.30 (22.63)	49.95 (23.32)
	Accuracy (Correct Word Sequences-Incorrect Word Sequences)	59.18 (26.50)	36.36 (20.79)
	Quality	3.67 (1.26)	2.44 (1.12)
	Written sentence generation	21.13 (8.63)	15.35 (8.22)
Process measure	Average writing burst length in words	6.02 (3.48)	3.82 (2.81)
Proximal factors			
Ability score	Spelling	20.39 (5.56)	15.90 (4.27)
Raw score	Handwriting fluency raw score	39.68 (16.44	26.82 (11.41)
Distal factors			
Raw scores	Listening span (WM)	2.40 (0.83)	1.74 (0.80)

	Oral vocabulary	143.51 (18.28)	121.44 (17.06)
	Oral 'sentence' raw	98.84 (19.66)	76.33 (15.23)
	Single word reading	95.23 (15.67)	97.15 (16.59)
Ability score	Reading comprehension	55.42 (19.46)	46.05 (7.67)