Assessing children's writing products using curriculum-based measures of writing (CBM-W)

Word Count:

With references: 7891

Without references: 6040

First Author: Will Piercy

Address:

6 Telford Close,

London,

SE19 3AG

Email: william.piercy@ucl.ac.uk

Second Author: Julie E. Dockrell

Address: 103

UCL Institute of Education

University College of London

25 Woburn Square

London

WC1H 0AL

Email: j.dockrell@ucl.ac.uk

We confirm that all authors approve the submission and that the paper is our original work and not under consideration elsewhere.

Assessing children's writing products using curriculum-based measures of writing (CBM-W)

Abstract

Aim: Writing is key to academic success and job opportunities later in life. Nevertheless, every year children leave primary school struggling to write and few assessments enable schools to reliably identify and target weaknesses in children's writing. This study explored curriculum-based measures of writing (CBM-W) for assessing writing in primary schools in the United Kingdom (UK).

Method: Twenty-seven children from a Year Five class in an English primary school responded to a narrative prompt administered for five minutes once a fortnight for a 10-week period. Children completed a standardised assessment of writing quality (the WIAT) prior to the first CBM-W assessment. CBM-W measures included number of words written (WW), words spelled correctly (WSC), correct word sequences (CWS; any two adjacent words used correctly in context) and correct minus incorrect word sequences (C-IWS).

Findings: All four CBM-W measures demonstrated concurrent validity in relation to the WIAT and captured significant differences in writing performance between children with and without a SEN. Measures also captured pupil growth, demonstrating the utility of CBM-W both as a tool for screening and progress monitoring.

Limitations: The small sample placed limits on the generalisability of findings. Additionally, piloting the equivalence of the writing prompts used in this study may have reduced the risk of measurement error.

Conclusion: Findings demonstrate the technical adequacy of CBM-W and promising evidence of its potential use by teachers, SENCOs and EPs to assess the writing of primary school children in the UK.

Keywords: curriculum-based measures of writing; CBM-W; writing assessment; primary school; progress monitoring.

Introduction

Writing is a complex skill which develops over time through interactions between the child's cognitive and linguistic systems, the instructional context, and the demands posed by the writing task (Kellogg, 2008). Writing skills are both a prerequisite for demonstrating and acquiring new knowledge at school (Dockrell et al., 2015) and key to academic achievement and employment opportunities later in life (Graham & Perin, 2007). Yet, many children are challenged by the writing process (Graham & Harris, 2004) and teachers in England report finding writing instruction difficult (Dockrell et al., 2016). This is reflected by recent figures showing that 22% of children in England left primary school below age related expectations for writing and disadvantaged children made significantly less progress than peers with similar starting points (DfE, 2019). To increase the overall number of children leaving primary school as proficient writers, timely and effective interventions are required for those with the greatest level of need.

Assessment plays a key role in the process of effective teaching and learning of writing (Jones, 2002), and can support teaching, both conceptually and practically (White, 1985). However, researchers have struggled to identify reliable and valid measures to assess children's writing (Meier et al., 2006; Rezaei & Lovorn, 2010) and research on writing has lagged behind research on reading and oral language development (Dockrell et al., 2015). Consequently, writing assessment has come to be viewed as problematic (Dockrell & Connelly, 2021), considered by some researchers to be one of the greatest obstacles to progress in writing instruction and research (Cole et al., 1997).

Nevertheless, over the past four decades promising findings have emerged from studies on curriculum-based measures (CBM; Deno, 1985), a commonly used method in North America for monitoring academic progress and making educational decisions in mathematics, reading and written expression (Lembke et al., 2013). Curriculum-based measures of writing (CBM-W) are likely to be attractive to teachers as they are quick to administer (typically between 3-7 minutes), involve relatively straightforward scoring procedures and are underpinned by a growing body of evidence to support their technical adequacy (e.g., reliability and validity) (McMaster & Espin, 2007; McMaster et al., 2011). Two recent meta-analyses (Romig et al., 2017; Romig et al., 2021) established the validity of a range of scoring procedures, prompt types and writing durations used in CBM-W assessments. Although the authors only established moderate criterion validity coefficients for CBM-W, they found this was similar to other published writing assessments and hypothesised that lower criterion validity should be expected for measures of written expression, given that writing is a complex and multi-faceted construct which is difficult to assess.

While findings from CBM-W research are promising, most studies have investigated these measures at single points in time (static) and there is a need for more research to examine the usefulness of these measures for tracking progress over time (slopes) and in classroom settings (Fuchs, 2004). As McMaster et al. (2017) highlight, progress monitoring plays a key role in enabling teachers to evaluate children's responsiveness to interventions and informing decisions about when instructional changes are needed.

Assessment is one of the core functions of the United Kingdom educational psychologist (EP) (Farrell et al., 2006), giving EPs an important role to play in helping schools to identify and support the learning needs of struggling writers. As Atkinson et al. (2022) point out, EPs use assessment not only for norm-referencing and enabling schools to access resources, but for formulating hypotheses about a child's strengths and difficulties, which can be tested through response to intervention. Additionally, given the recent surge in the testing expertise of non-EP school-based professionals, partly necessitated by the increasing onus on schools to organise access arrangements for English General Certificate of Secondary Education (Woods et al., 2018), EPs may also play an important role in training teachers and SENCOs to administer writing assessments themselves. This paper examines whether CBM-W are reliable and valid approaches to profile, monitor and support struggling writers in primary schools.

Theories of writing development and assessment

A key question driving research in the field of children's writing development is how children's writing should be assessed. One argument is that assessments should capture the skills children develop as they learn to write (Dockrell et al., 2015); this approach has the advantage of matching assessment to the developmental goals set for children (Dunsmuir et al., 2015). The 'Simple View of Writing' (SVW; Berninger et al., 2002) provides a framework for understanding developing writing skills as it captures both lower- and higher-level writing skills, differentiates across a range of ages and skill-levels, and can be used to identify the potential foci of writing instruction (Dockrell et al., 2017). The model represents writing development, figuratively, as a triangle consisting of three key components. At the base of the triangle, transcription skills (spelling and handwriting) and executive functions (including self-evaluation and goal-setting) enable the ultimate goal of text generation to occur (the process of translating ideas into meaningful segments of words, phrases and sentences).

Transcription skills account for the majority of variance in the writing quality of English speaking primary school children and are thought to constrain text production until they have become automatic (Graham et al., 1997). Specifically, having to attend to lower-level skills, such as searching for the letters to create words, can overload the capacity of working memory, leading ideas, plans and revising to be lost (Dockrell & Connelly, 2021). Additionally, children's handwriting (or typing) may not be fast enough to keep up with their thoughts, interfering with text generation and recall of ideas or text already planned and held in working memory (Graham et al., 1997). It is by developing a solid foundation in translation-level skills (i.e., text generation and transcription) that the writer is able to dedicate attentional resources, such as working memory, to higher level processes, such as organisation, revision, and planning (Wagner et al., 2011; Puranik et al., 2008). Therefore, to effectively evaluate the writing performance of struggling writers, measures should be sensitive to change in these basic skills.

Dimensions of the written text

The number of dimensions thought to underpin children's writing, and the measures which should be used to evaluate these, has been the subject of debate for some time (Dockrell & Connelly, 2021). Whilst earlier studies distinguish between productivity (i.e. the overall length of the written text, either in words or sentences) and text quality (i.e. the overall success with which the writer conveys meaning at discourse level) (Hall-Mills, 2009), recent research has focused on the dimensions of complexity and organisation (Puranik et al., 2008; Wagner et al., 2011). Complexity considers numbers of clauses, embedded clauses and more complex grammar (see, for example, Puranik, 2008). Organisation, by contrast, refers to the use of specific textual features to structure the written text (e.g. the use of topic sentences, logical ordering of ideas, and key elements (Wagner et al., 2011). Another approach with promising evidence has been to combine measures tapping written accuracy (e.g. spelling, punctuation and grammatical errors) and productivity (Dockrell et al., 2015; Allen et al., 2020). This captures both text generation and transcription components of the SVW and, as such, may prove useful for assessing the quality of written texts produced by developing writers.

Other factors also need to be considered when evaluating the utility of writing assessments. Research has demonstrated that using reliable and valid measures can improve children's spelling and writing when used for screening, monitoring purposes, and making instructional changes (Jung et al., 2017). Measures should also be sensitive to the performance of struggling writers and those with special educational needs (SEN) who generally produce shorter, less well organised texts that contain more mechanical, spelling and grammatical errors than typical writers (MacArthur et al., 2006). They should also capture growth in writing, particularly for struggling writers and those with SEN, who tend to progress at a slower rate and for whom traditional assessments may not capture progress (Deno, 2003); tracking progress over time is key for identifying changes in performance and the child's strengths and needs (Mansell et al., 2009; Qualifications and Curriculum Authority—QCA, 1999; Department for Education—DfE, 2011). Assessments should be quick to administer and use scoring procedures that facilitate communication between educational practitioners. Measures should capture children's writing performance across different genres since writers perform differently across writing genres (e.g. expository conveying facts and narratives describing situations) (Scott & Windsor, 2000). Finally, in line with the SVW (Graham et al., 1997), measures should capture the dimensions of productivity and accuracy in children's writing; specifically, how much text has been written and whether that text is accurate in form and meaning (Allen et al., 2020).

Assessment of written texts

Various approaches exist for evaluating children's written text products. Holistic ratings and rubrics involve giving a single score for the overall global quality of the written text based on descriptors developed by experts: e.g. Wechsler Individual Achievement Test III (WIAT; Pearson, 2009). These measures allow the assessor to give a single score with relatively little time involvement but do not reliably differentiate across writing levels, monitor change over time or provide information about children's performance on different writing dimensions (Espin et al., 2000). By contrast, analytic scoring involves using a range of scoring procedures to evaluate performance across different dimensions of writing. Scores can be examined separately to create profiles or combined to provide a single score (Huot, 1990). These measures can provide teachers with useful information about children's strengths and weaknesses in different aspects of writing and have demonstrated utility for examining writing trajectories of children with developmental difficulties (Connelly et al., 2006; Dockrell et al., 2007). However, analytic scoring procedures can be complicated, and inter-rater reliability is generally lower than for holistic measures (Dockrell & Connelly, 2021). CBM-W have been suggested as a promising alternative approach to writing assessment in primary school children.

Curriculum-based measures of writing

CBM-W were originally designed as a way of directly observing children's basic writing skills (Deno, 1985) and are a reliable and valid assessment tool for measuring writing proficiency in writers between the ages of seven and 12 (Weissenburger & Espin, 2005). CBM-W tasks involve children writing for short periods of time (typically between 3-7 minutes) in response to a written, spoken, or pictorial prompt, after which the writing is scored using measures designed to capture dimensions of productivity and accuracy (see Tables 1 and 2).

CBM-W have several benefits. They are quick to administer, reliably distinguish between writers of different ages and skill levels (Tindal & Parker, 1991) and can be used to identify struggling writers (McMaster et al., 2012; Dockrell et al., 2015). CBM-W tasks capture children's differential performance across genres (Dockrell et al., 2015) and monitor progress in children's writing over time (McMaster et al., 2011; McMaster et al., 2017).

Nevertheless, the use of CBM-W has not gone unchallenged. A key issue has been mixed findings in relation to the technical adequacy of CBM-W. Inter-rater reliability has been shown to vary ac-

Table 1

CBM-W measures used to capture written productivity

Measure (abbr.)	Measure	Definition
WW	Words written	The total number of words produced in the written text. Errors in spelling, word usage, capitalisation and punctuation are discounted (Fuchs & Fuchs, 2011). Numerals and crossed-out words are excluded from the final score.

CWS	Correct word sequences	Any two adjacent words in a sentence spelled and used correctly in context (Videen et al., 1982)

Table 2

CBM-W measures used to capture written accuracy

Measure (abbr.)	Measure	Definition
WSC	Words spelled correctly	The number of correctly spelled words, irrespective of proper usage, capitalisation and punctuation (Fuchs & Fuchs, 2011)
C-IWS	Correct minus incorrect word sequences	The number of correct word sequences minus the number of incorrect word sequences (Espin et al., 2008)

cording to the type of text produced, the age of the children studied and the scoring procedures under review (McMaster et al., 2012). Importantly, criterion validity coefficients (the extent to which children's CBM-W scores correlate with standardised assessments) have emerged as weak or non-significant, with the strongest coefficients ranging from between .35 and .50 in relation to holistic ratings of writing quality (Espin et al., 2000; Parker et al., 1991a).

Another issue is that few CBM-W studies have taken place outside of North America, therefore it is unclear whether CBM-W can be applied within other educational contexts where different teaching and learning practices are in operation (Dockrell et al., 2015). Finally, as most CBM-W research has focused on static scores, more work is needed to examine whether CBM-W can be used to monitor change in children's writing over time (McMaster et al., 2017). Monitoring change over time allows for the evaluation of instructional changes and the impact of interventions on writing development.

Progress Monitoring

There is a growing body of evidence to suggest CBM-W can be used for progress monitoring. Ritchey and Coker (2013) administered CBM-W narrative prompts and picture tasks to 170 children in Reception to Year Three across three separate time points and Dockrell et al. (2015) used narrative and expository prompts with 236 children in Years Three to Five at two separate time points. Although both studies captured growth using words written (WW), words spelled correctly (WSC) and correct word sequences (CWS) scores, to produce stable growth slopes, it is argued that data collection needs to take place on a more frequent basis (McMaster et al., 2017) and more data points are required (Christ et al., 2013). McMaster et al. (2011) addressed this by administering 12 weekly CBM-W tasks to 85 children in Year One and found that WW, WSC and CWS showed sensitivity to growth after 8 weeks. In one of the few studies to examine the use of CBM-W for monitoring progress in older primary school children, McMaster et al. (2017) administered weekly CBM-W tasks for a period of 12 weeks to children in Years 2-3 and 4-5, using CWS and correct minus incorrect word sequences (C-IWS). Steady linear growth was found for younger children across all 12 weeks, but growth began to taper off after 10 weeks for older children. Thus, further work is needed to establish the utility of CBM-W for monitoring writing progress in older primary school children.

Present study

The present study aims to extend our understanding of the utility of CBM-W to assess writing skills and development in a cohort of English primary school children. To achieve this objective four research questions were addressed:

1) To what extent do CBM-W measures demonstrate validity when benchmarked against a standardised assessment of writing quality (the WIAT III)?

2) Which, if any, CBM-W measures (WW, WSC, CWS and C-IWS) reliably distinguishes between typical and struggling writers within the same class?

3) Are CBM-W sensitive to change in children's writing over time?

4) To what extent are CBM-W sensitive to differences in the developmental growth patterns of typical and struggling writers?

Methodology

Participants

A total of 27 children from one Year Five class were invited to participate in the study (boys *n* = 14; girls *n* = 13). There were seven children with a SEN and 20 children without an identified SEN as recorded in the class register. The class were selected from an inner London school and representative of the local authority for percentage of children on free school meals (FSM) and ethnicity but slightly above average in terms of those with a reported SEN. Ten pupils were eligible for FSM and two had English as an additional language. According to 2019 data, the school scored in line with the national average in Key Stage 2 writing assessments (progress score = 0.3, where 0 denotes expected progress; CI = -1.8 to 2.4).

Design

A repeated measures design was used where children completed one CBM-W task (narrative prompt) every fortnight over a period of 10 weeks during the summer term. A standardised measure of writing quality was administered at the beginning of the study. All CBM-W assessments were administered on the same day and time every two weeks, and a total of five narrative samples to different prompts were collected for each pupil.

Materials

Prompts were sourced from Intervention Central, an online learning resource for Response to Intervention (RTI) tools and resources (https://www.interventioncentral.org/teacher-resources/curriculum-based-measurement-probes-writing). To ensure prompts tapped similar levels of knowledge and skill, these were chosen by the teacher (the first author) based on a judgement of the interests and knowledge of children in the class. Children were given a lined sheet of paper with the prompt printed at the top of the page. Appendix A provides details of the prompts used.

Measures

Responses were scored on measures of productivity (WW and CWS) and accuracy (WSC and C-IWS). A complete list of measures and definitions is provided in Appendix B. The Written Expression subtest from the WIAT-III (Pearson, 2009) was completed by all pupils at the beginning of the study. The test has five sections, three of which can be used with children aged 8 to 11 years: Word Fluency, Sentences and Paragraph. The paragraph writing task was chosen as it was appropriate for use with the participants (aged 9-10 years) and given its similarity to the CBM-W tasks administered; specifically, both tests involve children producing a written response to a prompt in a relatively short period of time (10 mins for the WIAT and 5 mins for CBM-W). As the original prompt provided in the WIAT-III was considered more likely to elicit expository writing from children (e.g. 'My favourite game is…'), this was substituted for a narrative prompt (e.g. 'The zookeeper noticed that the cage was open and…') to ensure consistency with those used in the five CBM-W tasks administered in the study. Children's written composition can be scored on the WIAT either holistically or analytically: reliability = .94, correlation with Woodcock-Johnson Psycho-Educational Battery-Revised, Broad Written Language Cluster = .70. The analytic scoring rubric was used to assess children's written composition in relation to UK standardised norms. This included three evaluation categories or dimensions: Mechanics (i.e. spelling and punctuation errors), Organisation (i.e. sentence structure, paragraph length and linking expressions), and Vocabulary (i.e. lexical diversity and interesting expressions), the maximum scores for which are 9, 10 and 5 points, respectively. Children's scores on each of these categories are combined to provide a single score out of 24 points which can be standardised against UK normative data.

Procedure

At the beginning of each CBM-W assessment, the first line of the story was read to children, and they were given 30 seconds to plan and five minutes to write the rest of the story to the best of their ability. Children were encouraged to do their best with spelling and at the end of the test were instructed to stop writing and put a dash after the last word they had written to ensure only writing produced within the five minutes would be scored.

Ethical approval was obtained from University College of London (UCL) which follows British Psychological Society (BPS) guidelines. Both children and the school were given information sheets outlining the purpose and aims of the study before being asked to give written informed consent. A child-friendly information sheet was provided which the class teacher read through with children, explaining their right not to participate and withdraw from the study at any point. Children were given a consent form which they signed if they were willing to participate. Parents were given an information sheet and asked to opt out by a given date should they wish their child to be excluded from the study. Children completed tests as a whole class and the school's Covid-19 policy was carefully adhered to throughout the study.

Coding and reliability

Children's CBM-W writing samples were photocopied and scored by the main researcher. Inter-rater reliability was checked for 10% of all scripts. Three writing samples from each of the five assessment points were randomly selected and marked by a second assessor, resulting in a total of 15 writing samples being checked for reliability. These were scored by another teacher at the children's school who was given training on using scoring procedures. Cronbach's Alpha was above 0.8 for all measures indicating strong reliability across both assessors.

Results

CBM-W measures and standardised measures of writing

To examine the concurrent validity of CBM-W, correlations between children's WW, CWS, WSC and C-IWS scores at the first test point and their raw scores on the WIAT measure were examined. Correlations are presented in Table 3. Pearson's *r* was used on account of the fact both variables were continuous and normally distributed (Shapiro-Wilk's, $p > .05$). Children's mean scores in the WIAT were at the lower end of the average range ($M = 91.50$, SD = 4.68). As Table 3 shows there

Table 3

Pearson correlations between CBM-W narrative performance and WIAT raw scores

Writing measure	**1**	**2**	**3**	**4**
1. WIAT Raw score				
2. Words Written	0.57**			
3. Correct word sequences	0.73**	0.95		
4. Words spelled correctly	0.69**	0.99	0.95	
5. Correct minus incorrect word sequences	0.78**	0.88	0.98	0.89

**. Correlation is significant at the 0.01 level (2-tailed).

were significant and strong correlations between scores on all CBM-W measures and children's total WIAT total score, demonstrating that measures designed to capture productivity and accuracy in children's writing served as a valid indicator of children's overall writing proficiency. Of all four measures, CWS and C-IWS demonstrated the strongest correlations with the WIAT.

CBM-W measures and sensitivity to differences in children's writing levels

A total of 27 children contributed narrative texts in the study. As shown in Tables 4 and 5, large standard deviations for each of the four measures across all five time points indicated a high degree of variability amongst children's writing scores. To further explore the extent to which CBM-W were sensitive to differences in children's writing levels, independent-samples t-tests were conducted to determine whether there were any differences in mean scores between children with and without SEN for each measure at the first assessment point. There were seven children with SEN and 20 children without SEN. Two outliers were identified within the WW data and one within the CWS data, as assessed by inspection of a box plot for values greater than 1.5 box-lengths from the edge of the box. Analyses were run with and without inclusion of the outliers. Scores for all measures were normally distributed (Shapiro-Wilk's, $p > .05$), and there was homogeneity of variances (Levene's test for equality of variances WW, $p = .563$; WSC, $p = .563$; CWS, $p = .179$; C-IWS, $p = .090$). Children with no SEN performed significantly higher on all measures than those with SEN: WW (SEN M = 37.14, SD = 16.58; No SEN M = 69.85, SD = 20.89), $t(25) = 3.73$, $p < .05$; WSC (SEN M = 28.14, SD = 12.25; No SEN M = 66.75, SD = 21.06), $t(25) = 4.55$, $p < .05$; CWS (SEN M = 24.43, SD = 11.93; No SEN M = 65.55, SD = 22.30), t(25) = 4.61, p < .05; C-IWS (SEN M = 15.43, SD = 10.77; No SEN M = 56.10, SD = 23.24), $t(25) = 4.42$, $p < .05$. Children with SEN were performing significantly lower than those without a SEN on all measures and these differences were supported when children's performance was compared using an independent samples t-test on the standardised writing measure (SEN M = 8.00, SD = 1.63; No SEN M = 16.50, SD = 3.07), $t(25) = 6.930$, $p < .05$.

CBM-W measures and sensitivity to change over time

A two-way mixed ANOVA was conducted to explore the sensitivity of CBM-W to change in children's writing across the 10-week observation period, and to determine whether there were any significant differences in how children with and without a SEN developed in each of the CBM-W measures over time. Tables 4 and 5 show pupils' productivity and accuracy scores across the five assessment points.

At the first assessment point, the outliers found in WW data and 1 outlier in CWS were removed. Data were normally distributed (Shapiro-Wilk test (p > .05) and the assumption of sphericity was not violated (Mauchly's test of sphericity [WW $\chi2$ (9) = 8.95, p = .44; CWS $\chi2$ (9) = 5.99, p = .74; WSC $\chi2$ (9) = 10.68, p = .30; C-IWS $\chi2$ (9) = 8.56, p = .48].

Table 4

Mean (standard deviation) productivity scores for narrative probes for year 5

Time point	Words written		Correct word sequences	
	M	*SD*	*M*	*SD*
1	61.37	24.41	54.89	27.08
2	71.93	25.24	64.52	25.33
3	69.41	26.03	61.56	26.59
4	69.00	23.57	64.26	24.84
5	66.63	28.11	61.15	28.86

Table 5

Mean (standard deviation) accuracy scores for narrative probes for year 5

Time point	Words spelled correctly	Correct minus incorrect word sequences

	M	*SD*	*M*	*SD*
1	56.74	25.61	45.56	27.41
2	67.41	26.29	51.22	26.51
3	64.93	26.91	50.11	28.62
4	65.70	24.24	54.22	26.96
5	62.48	27.88	51.07	30.81

There were no statistically significant interactions between time and SEN status for any of the measures [WW F (4, 92) = 0.71, p = .59, $\eta 2p$ = .124; CWS F (4, 96) = 1.19, p = .32, $\eta 2p$ = .047; WSC F (4, 100) = 0.69, p = .60, $\eta 2p$ = .153; C-IWS F (4, 100) = 1.06, p = .38, $\eta 2p$ = .096]. According to Cohen's (1988) criteria, where partial eta squared ($\eta 2p$) values of .01, .06, and .14 indicate small, medium and large effect sizes, effect sizes found here were in the medium range. Line graphs in Figures 1 and 2 present the writing trajectories of children with and without a reported SEN. The main effect of time showed statistically significant differences for all CBM-W measures [WW F (4, 92) = 3.25, p < .05, $\eta 2p$ = .12; CWS F (4, 96) = 5.20, p < .05, $\eta 2p$ = .18; WSC F (4, 100) = 4.53, p < .05, $\eta 2p$ = .15; C-IWS F (4, 100) = 2.65, p < .05, $\eta 2p$ = .10], with medium to large effect sizes (Cohen, 1988). Post hoc analysis with a Bonferroni adjustment revealed no statistically significant increases between any of the five time points for WW. However, statistically significant increases were found: for CWS, from time point 1 to 2 [9.63 (95% CI, 1.07 to 19.54), p < .05] and time point 1 to 4 [9.37 (95% CI, 1.85 to 17.30), p < .05]; for WSC, from time point 1 to 2 [10.67 (95% CI, 1.48 to 19.85), p < .05] and time point 1 to 4 [8.96 (95% CI, 1.14 to 16.79), p < .05]; and for C-IWS, from time point 1 to 4 [8.67 (95% CI, 0.47 to 16.86), p < .05].

The main effect of group indicated that there were statistically significant differences in mean scores between SEN and No-SEN groups for all CBM-W measures [WW F(1, 23) = 8.73, p < .05, $\eta 2p$ =

.28; CWS F (1, 24) = 22.00, p < .05, $\eta 2p$ = .48; WSC F (1, 25) = 17.33, p < .05, $\eta 2p$ = .41; C-IWS F (1, 25) = 21.73, p < .05, $\eta 2p$ = .47], with large effect sizes (Cohen, 1988). A comparison of the writing performances of typical and struggling writers in all CBM-W measures is presented in Figure 1 and 2.

Figure 1

A comparison of developmental growth patterns in written productivity (WW = words written: CWS = correct word sequences) for pupils with and without SEN

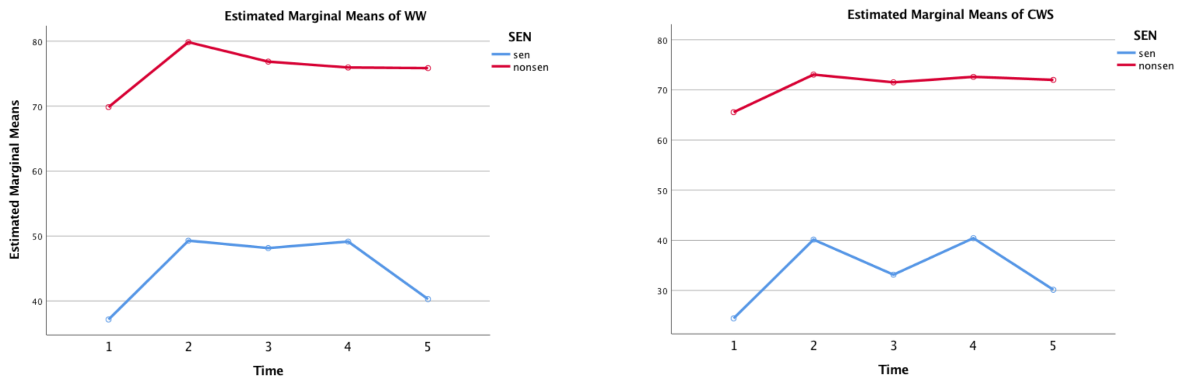
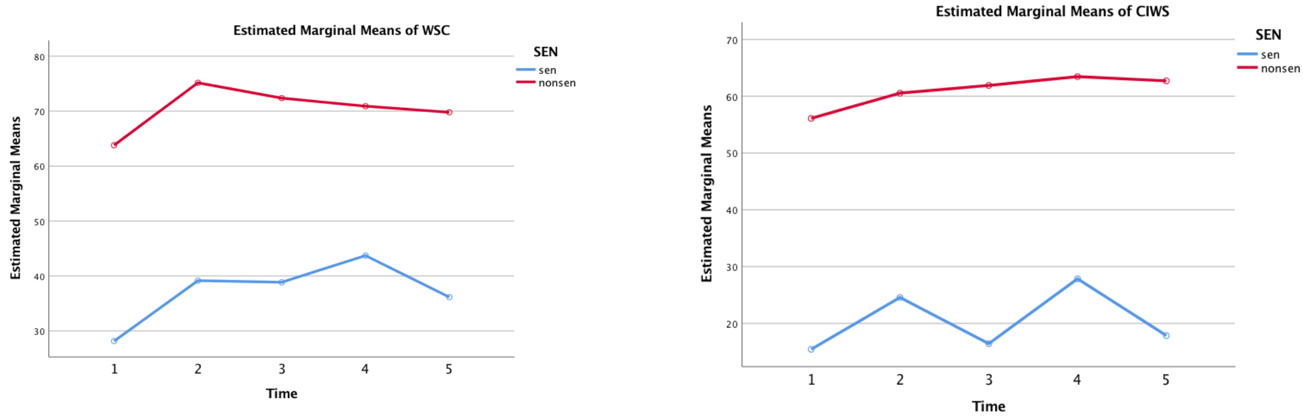


Figure 2

A comparison of developmental growth patterns in written accuracy (WSC = words spelled correctly: C-IWS = correct minus incorrect word sequences) for pupils with and without SEN



Discussion

The study aimed to examine the utility of CBM-W for assessing writing in primary schools in the UK. We first aimed to establish the technical accuracy of CBM-W in our sample (Weissenburger & Espin, 2005; Tindal & Parker, 1991). High inter-rater reliability across assessors and strong correlations between all four CBM-W measures and the WIAT demonstrated that measures are capturing pupils' written accuracy and productivity and can serve as a valid marker of writing proficiency in upper primary school children.

Secondly, we aimed to examine whether the CBM-W measures captured variability in writing performance. All four CBM-W measures differentiated across writers from the same class, as indicated by the high degree of variability in scores during the 10 weeks of observation. The marked heterogeneity in scores for all CBM-W measures was reflected in children's WIAT scores. These results are consistent with national data on writing tests for England, which also show a large amount of variability within year groups (DfE, 2011, 2012). Large and significant differences were found on both the WIAT and the CBM-W measures for children with and without SEN, further confirming the use of the measure for capturing struggling writers (Dockrell et al., 2015; McMaster et al., 2017).,

Sensitivity to change

In relation to the third research question, we found that three of the four measures (CWS, WSC and C-IWS) captured change in children's writing over time, reflecting both dimensions of productivity and accuracy in children's writing. An examination of children's writing trajectories on each measure over time revealed a common trend where for all measures children made steep progress between weeks one to three which began to taper off thereafter. Data from the second assessment point were re-examined to explore potential causes for the spike in children's scores at week three; this confirmed a normal distribution of the data with no extreme outliers that may have occurred as the result of an error in scoring or the influence of one or two children. One explanation is that the narrative prompt used in the second assessment was either more accessible or relevant to children, thereby tapping greater levels of knowledge and skill. An alternative explanation for this deceleration in progress is that there was a general decline in children's motivation to respond to CBM-W prompts over time. Given that testing took place in the summer term, seasonal effects may have come into play (Christ et al., 2010; Fuchs et al., 1993). A similar trend was observed by McMaster et al. (2017), who found that Year 4-5 children's CWS and C-IWS scores started to taper off after 10 weeks whereas younger children (Year 2-3) continued to make progress across the whole 12-week monitoring period.

Developmental differences in CBM-W performance are worthy of further research. As children become older and more skilled as writers, attending more closely to higher level aspects of writing (e.g. use of vocabulary, structure, genre and the development of new and complex ideas), CBM-W may not capture gains in children's writing proficiency (McCutchen, 2006). Whilst in the current study there was some indication that children's writing performance began to level off for WW and WSC, this does not appear to have been the case for children's CWS and C-IWS scores, which continued to show some linear growth after week three. One possibility for these trends is that CWS and C-IWS, which capture more advanced writing skills (e.g. writing complexity, word form and use, and errors), are more appropriate for use with older children than simpler measures (i.e. WW and WSC) which tap into basic skills (e.g. spelling accuracy and the amount written; Allen et al., 2020). This may also explain why CWS and C-IWS had the strongest concurrent validity in relation to the WIAT; specifically, by Year Five, measures which tap more advanced writing skills, may provide a more accurate indicator of children's writing proficiency and progress over time.

Growth patterns of typical and struggling writers

Our final research question explored whether CBM-W were sensitive to differences in the developmental growth patterns of typical and struggling writers. Results from the mixed two-way ANOVA indicated no significant interactions between SEN status and change in writing over time for any of the CBM-W measures. However, the small sample used likely provides insufficient power to detect differences. Indeed, Allen et al. (2020) found that children in Years 2-5 who started lower on CWS and C-IWS made greater progress than those with higher initial scores. Therefore, writing trajectories of children with and without a reported SEN were compared by visual inspection of line graphs (see figures 1 and 2) to establish whether any differences in developmental trends could be discerned.

An inspection of the graphs as shown in Figures 1 and 2 revealed that, while for children without a reported SEN C-IWS scores continued to gradually increase over time, there was a trend for progress to taper off after the second assessment point for all other measures. By comparison, an examination of the writing trajectories of children with SEN provided a more mixed picture. Whilst overall children with SEN appeared to make some linear progress in measures of accuracy (WSC and C-IWS), in measures of productivity (i.e. WW and CWS) progress appeared to wane over time. It may be that children with SEN attended more closely to the accuracy of their writing and, as such, their written productivity became constrained. Additionally, compared with typical writers, the writing trajectories of the children with SEN tended to be less stable. By the end of Year Five, gains made by struggling writers in lower-level skills (e.g. handwriting and spelling) may still be captured whereas for their typically developing peers these measures are no longer as sensitive to growth. Another possibility is that the prompts were not equivalent and elicited differences in the subject knowledge of children with and without a SEN; for instance, being asked to write a story about an astronomer in the third assessment may have placed a greater burden on the executive function of struggling writers less fluent in basic writing skills than for typical writers, resulting in the production of less grammatically accurate texts and, accordingly, lower CWS and C-IWS scores for this week. This is consistent with the findings of McMaster et al. (2017) who attributed similar variations in the CWS and C-IWS scores of children in Years 2-5 to prompt-specific differences (e.g. children’s interest or ease in responding to certain prompts).

Limitations and directions of future research

The findings of this study need to be considered in light of various limitations. First, writing data were collected from a small sample of children over a relatively short period of time, limiting the generalisability of findings to the wider population. Additionally, more data points in the study would have helped to capture children's writing over time and understand changes in the developmental growth patterns of typical and struggling writers. Research has shown that as data points are added children's writing slopes become more stable and reliable (McMaster et al., 2011) and there is a lower risk of measurement error (Hintze & Christ, 2004). Additionally, piloting the equivalence of the writing prompts used in this study would have reduced the risk of measurement error. Given evidence that writers perform differently across writing genres (Scott & Windsor, 2000), it is also possible that by using a narrative prompt instead of the expository prompt provided in the WIAT-III, the scoring norms for this measure were invalidated; nonetheless, there is insufficient research examining primary school children's writing across different genres. Seasonal effects may have acted as a confound in this study; specifically, as children's writing performance was observed during the summer term, change in specific CBM-W indices of progress may have been influenced by shifts in motivation or instruction rather than signifying the appropriateness of these measures for children in Year Five. Future research trialling CBM-W across an entire school year, and collecting information about instructional foci, could elucidate the role of teacher effects and children's writing performance at different times of year (McMaster et al., 2017). It may also prove advantageous to explore the impact of other factors likely to influence children's progress in CBM-W (e.g. the difficulty of writing prompts, the age and skill of writers, and children's self-regulation and working memory). Finally, future work should aim to collect more information about children to supplement CBM-W data (e.g. type of difficulty, reading levels, spelling and handwriting, language status and FSM eligibility), which have been shown to be related to children's writing proficiency (Dockrell et al., 2019). Doing so, may help to explore potential causes for poorer performance in specific CBM-W measures, as well as providing a further source of evidence for the criterion validity of the measure.

Implications of findings for professionals

The findings of this research have several implications for schools and EPs in terms of how they might utilise CBM-W to identify and support struggling writers in a timely manner.

1. At the individual level, CBM-W might be used by EPs to profile a child's developing writing skills, information from which might be used by teachers, alongside the EP, to jointly identify areas of strength and weakness, and the type of support needed.

2. At the group level, CBM-W could be used by teachers and EPs to identify the foci of writing interventions for groups of learners struggling with similar aspects of writing, thereby having the potential to support the needs of greater numbers of children.

3. At the whole school level, EPs could train teachers and SENCOs to administer and score CBM-W assessments themselves. Training teachers to use a writing assessment with evidence of validity and reliability has the potential to promote equity of access to timely and effective intervention for children with the greatest level of need.

4. As well as using these assessments to identify struggling writers and the specific areas for targeted support, CBM-W could also be used to track change in children's writing performance over time, thereby monitoring the impact of any interventions used on their developing writing skills.

Conclusion

In summary, writing is a key skill which children need to learn to succeed in school and later life, but many children struggle to become proficient writers. Therefore, teachers need access to reliable and valid assessments for supporting progress in the writing skills of struggling writers. Findings from this study suggest that CBM-W may serve as a useful, and quick to administer, assessment tool for teachers and SENCOs, alongside EPs, to evaluate the written texts of children in later primary school. Although further research is needed to understand which measures are most appropriate for children of different ages and levels of skill, this is one of few studies to examine the utility of CBM-W for assessing primary school children in the UK (Dockrell et al., 2015), and provides promising evidence that CBM-W can be used to monitor and support progress in children's writing.

References

Allen, A.A., Jung, P.-G., Poch, A. L. et al. (2020). Technical Adequacy of Curriculum-Based Measures in Writing in Grades 1-3. *Reading & Writing Quarterly*, *36*(6), 563–587.

Atkinson, C., Barrow, J., & Norris, S. (2022). Assessment practices of educational psychologists and other educational professionals. *Educational Psychology in Practice, 38*(4), 347–363.

Berninger, V.W., Vaughan, K., Abbott, R. D. et al. (2002). Teaching Spelling and Composition Alone and Together: Implications for the Simple View of Writing. *Journal of Educational Psychology*, *94*(2), 291–304.

Birch, S., Gulliford, A. & Cline, T. (2015). *Educational Psychology*. Taylor and Francis.

Christ, T.J., Monaghen, B. D., Zopluoglu, C. et al. (2013). Curriculum-Based Measurement of Oral Reading: Evaluation of Growth Estimates Derived With Pre–Post Assessment Methods. *Assessment for Effective Intervention*, *38*(3), 139–153.

Christ, T.J., Silberglitt, B., Yeo, S. et al. (2010). Curriculum-Based Measurement of Oral Reading: An Evaluation of Growth Rates and Seasonal Effects Among Students Served in General and Special Education. *School Psychology Review*, *39*(3), 447–462.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge.

Coker, D.L., & Ritchey, K. D. (2010). Curriculum-Based Measurement of Writing in Kindergarten and First Grade: An Investigation of Production and Qualitative Scores. *Exceptional Children*, *76*(2), 175–193.

Cole, J., Haley, K., & Muenz, T. (1997). Written expression reviewed. *Research in the Schools,* 4(1), 17–34.

Connelly, V., Campbell, S., MacLean, M. et al. (2006). Contribution of Lower Order Skills to the Written Composition of College Students With and Without Dyslexia. *Developmental Neuropsychology*, *29*(1), 175–196.

Deno, S. (1985). Curriculum-Based Measurement: The Emerging Alternative. *Exceptional Children*, *52*(3), 219–232.

Deno, S. (2003). Developments in Curriculum-Based Measurement. *The Journal of Special Education*, *37*(3), 184–192.

Department for Education. (2011). The Framework for the National Curriculum: a Report by the Expert Panel for the National Curriculum Review. London: DfE [online]. Available online at: https://www.gov.uk/government/uploads/system/uploads attachment_data/file/175439/NCR-Expert_Panel_Report. pdf (accessed 18 September 2013).

Department for Education. (2012). National Curriculum assessments. Available online at: https://www.gov.uk/government/collections/statistics-key-stage-2(accessed 18 September 2013).

Department for Education. (2019). National curriculum assessments at key stage 2 in England, 2019 (provisional). London: DfE [online]. Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/830285/KS2_Provisional_publication_text_2019.pdf (accessed 11 September 2021).

Dockrell, J.E., Connelly, V., & Arfè, B. (2019). Struggling writers in elementary school: Capturing drivers of performance. *Learning and Instruction*, *60*, 75–84.

Dockrell, J. E., & Connelly, V. (2021). Capturing the challenges in assessing writing: Development and writing dimensions. In T. Limpo & T. Olive (Eds.), *Executive functions and writing* (pp. 103–135). Oxford University Press.

Dockrell, J.E., Connelly, V., Walter, K. et al. (2015). Assessing children's writing products: the role of curriculum based measures. *British Educational Research Journal*, *41*(4), 575–595.

Dockrell, J.E., Connelly, V., Walter, K. et al. (2017). The Role of Curriculum Based Measures in Assessing Writing Products. In *Writing Development in Struggling Learners* (Vol. 35, pp. 182–197).

Dockrell, J.E., Lindsay, G., Connelly, V. et al. (2007). Constraints in the Production of Written Text in Children with Specific Language Impairments. *Exceptional Children*, *73*(2), 147–164.

Dockrell, J.E., Marshall, C. R., & Wyse, D. (2016). Teachers' reported practices for teaching writing in England. *Reading & Writing*, *29*(3), 409–434.

Dunsmuir, S., Kyriacou, M., Batuwitage, S. et al. (2015). An evaluation of the Writing Assessment Measure (WAM) for children's narrative writing. *Assessing Writing*, *23*, 1–18.

Espin., C.A.(1999). Criterion-related validity of curriculum-based measures in writing for secondary school students. *Reading & Writing Quarterly*, *15*(1), 5–27.

Espin, C.A., Shin, J., Deno, S. et al. (2000). Identifying indicators of written expression proficiency for middle school students. *The Journal of Special Education*, 34(3), 140–153.

Espin, C.A., Wallace, T., Campbell, H. et al. (2008). Curriculum-Based Measurement in Writing: Predicting the Success of High-School Students on State Standards Tests. *Exceptional Children*, *74*(2), 174–193.

Farrell., P., Woods, K., Lewis, S. et al. (2006). *A review of the functions and contribution of educational psychologists in England and Wales in light of "Every child matters: change for children"*. Department for Education and Skills.

Fuchs, L.S. (2004). The Past, Present, and Future of Curriculum-Based Measurement Research. *School Psychology Review*, *33*(2), 188–192.

Fuchs, L.S, & Fuchs, D. (2009). On the Importance of a Unified Model of Responsiveness to Intervention. *Child Development Perspectives*, *3*(1), 41–43.

Fuchs, L. S., & Fuchs, D. (2011). Using CBM for progress monitoring in written expression and spelling. *US Office of Special Education*.

Fuchs, L.S., Fuchs, D., Hamlett, C.L. et al. (1993) Formative Evaluation of Academic Progress: How Much Growth Can We Expect? *School Psychology Review*, *22:1*, *27-48*.

Gansle, K.A., Noell, G. H., VanDerHeyden, A. M. et al. (2002). Moving Beyond Total Words Written: The Reliability, Criterion Validity, and Time Cost of Alternate Measures for Curriculum-Based Measurement in Writing. *School Psychology Review*, *31*(4), 477–497.

Graham, Berninger, V. W., Abbott, R. D. et al. (1997). Role of Mechanics in Composing of Elementary School Students: A New Methodological Approach. *Journal of Educational Psychology*, *89*(1), 170–182.

Graham, S., & Harris, K. (2004). Writing instruction, in: B. Wong (Ed.) Learning about learning disabilities (3rd edn) (San Diego, CA, Elsevier), 281–313.

Graham, S., Morphy, P., Harris, K. R. et al. (2008). Teaching Spelling in the Primary Grades: A National Survey of Instructional Practices and Adaptations. *American Educational Research Journal*, *45*(3), 796–825.

Graham, S & Perin, D. (2007). A Meta-Analysis of Writing Instruction for Adolescent Students. *Journal of Educational Psychology*, *99*(3), 445–476.

Hall-Mills, S. (2009). Linguistic Feature Development in Elementary Writing: Analysis of Microstructure and Macrostructure Features in a Narrative and an Expository Genre. *ProQuest Dissertations Publishing*.

Hempenstall, K. (2020). Teaching reading through Direct Instruction: A role for educational psychologists? *The Educational and Developmental Psychologist*, *37*(2), 133–139.

Hintze, J.M., & Christ, T. J. (2004). An Examination of Variability as a Function of Passage Variance in CBM Progress Monitoring. *School Psychology Review*, *33*(2), 204–217.

Huot, B. (1990). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research*, *60*(2), 237–263.

Jones, D. (2002). Keeping track: Assessment in writing. In M. Williams (Ed.), *Unlocking writing: A guide for teachers*, (pp. 92–105). London: David Fulton Publishers Ltd.

Jung, P-G., McMaster, K. L., & delMas, R. C. (2017). Effects of Early Writing Intervention Delivered Within a Data-Based Instruction Framework. *Exceptional Children*, *83*(3), 281–297.

Kellogg, R.T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, *1*(1), 1–26.

Lembke, E., Hampton, D., & Hendriker, E. (2013). Data-based decision-making in academics using curriculum-based measurement. In J. W. Lloyd, T. J. Landrum, B. Cook, & M. Tankersley (Eds.), *Research-based approaches for assessment* (pp. 18–31). Pearson.

MacArthur, C., Graham, S. & Fitzgerald, J. (2006) *Handbook of research on writing* (New York, Guildford).

Mansell, W., James, M. & Advisory Group. (2009). *Assessment in schools. Fit for purpose? A commentary by the Teaching and Learning Research Programme* (London, Economic and Social Research Council, Teaching and Learning Research Programme).

McCutchen, D. (2006). Cognitive Factors in the Development of Children's Writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 115–130). The Guilford Press.

McMaster, K.L., Du, X., Yeo, S. et al. (2011). Curriculum-Based Measures of Beginning Writing: Technical Features of the Slope. *Exceptional Children*, *77*(2), 185–206.

McMaster, K.L., & Espin, C. (2007). Technical Features of Curriculum-Based Measurement in Writing: A Literature Review. *The Journal of Special Education*, *41*(2), 68–84.

McMaster, K.L., Parker, D., & Jung, P.-G. (2012). Using Curriculum-Based Measurement for Beginning Writers Within a Response to Intervention Framework. *Reading Psychology*, *33*(1-2), 190–216.

McMaster, K.L., Ritchey, K. D., & Lembke, E. (2011). Curriculum-based measurement for beginning writers: recent developments and future directions. In *Assessment and Intervention* (Vol. 24, pp. 111–148). Emerald Group Publishing Limited.

McMaster, K.L., Shin, J., Espin, C. A. et al. (2017). Monitoring elementary students' writing progress using curriculum-based measures: grade and gender differences. *Reading & Writing*, *30*(9), 2069–2091.

McMaster, K.L., Xiaoqing Du, & Pétursdóttir, A.-L. (2009). Technical Features of Curriculum-Based Measures for Beginning Writers. *Journal of Learning Disabilities*, *42*(1), 41–60.

Meier, S.L., Rich, B. S., & Cady, J. (2006). Teachers' use of rubrics to score non-traditional tasks: factors related to discrepancies in scoring. *Assessment in Education: Principles, Policy & Practice*, *13*(1), 69–95.

Olinghouse, N.G. (2008). Student- and instruction-level predictors of narrative writing in third-grade students. *Reading & Writing*, *21*(1-2), 3–26.

Parker, R., Tindal, G., & Hasbrouck, J. (1991). Countable indices of writing quality: Their suitability for screening-eligibility decisions. *Exceptionality: The Official Journal of the Division for Research of the Council for Exceptional Children*, *2*(1), 1–17.

Parker, R., Tindal, G., & Hasbrouck, J. (1991). Progress Monitoring with Objective Measures of Writing Performance for Students with Mild Disabilities. *Exceptional Children*, *58*(1), 61–73.

Puranik, C.S., Lombardino, L. J., & Altmann, L. J. P. (2008). Assessing the Microstructure of Written Language Using a Retelling Paradigm. *American Journal of Speech-Language Pathology*, *17*(2), 107–120.

QCA (2008). Assessing pupils' progress: assessment at the heart of learning. Report for Qualifications and Curriculum Authority, Carrickfergus, UK.

Rezaei, A.R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, *15*(1), 18–39.

Ritchey, K.D., & Coker, D. L. (2013). An Investigation of the Validity and Utility of Two Curriculum-Based Measurement Writing Tasks. *Reading & Writing Quarterly*, *29*(1), 89–119.

Romig, J.E., Miller, A. A., Therrien, W. J. et al. (2021). Meta-Analysis of Prompt and Duration for Curriculum-Based Measurement of Written Language. *Exceptionality*, *29*(2), 133–149.

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

Scott, C.M., & Windsor, J. (2000). General Language Performance Measures in Spoken and Written Narrative and Expository Discourse of School-Age Children With Language Learning Disabilities. *Journal of Speech, Language, and Hearing Research*, *43*(2), 324–339.

Tindal, G., & Parker, R. (1989). Assessment of Written Expression for Students in Compensatory and Special Education Programs. *The Journal of Special Education*, *23*(2), 169–183.

Tindal, G., & Parker, R. (1989). Development of Written Retell as a Curriculum-Based Measure in Secondary Programs, *School Psychology Review, 18:3*, 328-343.

Tindal, G., & Parker, R. (1991). Identifying measures for evaluating written expression. *Learning Disabilities Research & Practice*, *58(1)*, 61–73.

Videen, J., Marston, D., & Deno, S. L. (1982). Correct word sequences: A valid indicator of proficiency in written expression (Vol. IRLD-RR-84). Minneapolis: MN: University of Minnesota, Institute for Research on Learning Disabilities.

Wagner, R.K., Puranik, C. S., Foorman, B. et al. (2011). Modeling the development of written language. *Reading & Writing*, *24*(2), 203–220.

Wagner, R.K., Smith, A., Allen, A. et al. (2019). Exploration of New Complexity Metrics for Curriculum-Based Measures of Writing. *Assessment for Effective Intervention*, *44*(4), 256–266.

Weissenburger, J.W., & Espin, C. A. (2005). Curriculum-based measures of writing across grade levels. *Journal of School Psychology*, *43*(2), 153–169.

Weschler, D. (2005). *Weschler Individual Achievement Test (WIAT-II UK)* (London, Pearson, Harcourt Assessments).

White, E. M. (1985). Teaching and assessing writing. San Francisco: Jossey-Bass.

Woods, K., James, A., & Hipkiss, A. (2018). Best practice in access arrangements made for England's General Certificates of Secondary Education (GCSEs): where are we 10 years on? *British Journal of Special Education*, *45*(3), 236–255.

Wright, J. *Written Expression Probe Generator. 'Intervention Central'.* https://www.interventioncentral.org/teacher-resources/curriculum-based-measurement-probes-writing.

Appendices

Appendix 1.

Narrative Prompts Responded to by Pupils

	Narrative Probes
WIAT	The zookeeper noticed that the cage was open and…
CBM-W 1	In the morning, I opened my door and saw five horses standing in the street. Then…
CBM-W 2	The girl found a toy in the yard that…
CBM-W 3	The astronomer was looking through the telescope at the moon when she noticed that…
CBM-W 4	Last week, a dog wandered into my school and…
CBM-W 5	When the snowstorm began, the lights went out just before…

Appendix B.

CBM-W measures used to assess pupils' written productivity and accuracy

Measure (abbr.)	Measure	Definition
WW	Words written	This was calculated by counting the total number of words produced in the 5-minute period. Errors in spelling, word usage, capitalisation and punctuation were not penalised (Fuchs & Fuchs, 2011) and numerals and crossed-out words were excluded from the final score.
WSC	Words spelled correctly	This was calculated by counting all correctly spelled words, irrespective of proper usage, capitalisation and punctuation (Fuchs et al., 2011)

CWS	Correct word sequences	This was calculated by counting any two adjacent words that were acceptable within the context of the sample to a native English speaker (Espin et al., 2005). To be considered ‘acceptable’, word sequences had to be syntactically and semantically correct (Videen et al.,1982). End punctuation and beginning capitalisation were taken into account (Tindal & Parker, 1989). To distinguish between transcription errors and text generation, spelling mistakes were not penalised (Dockrell et al., 2015)
C-IWS	Correct minus incorrect word sequences	This was the number of correct word sequences minus the number of incorrect word sequences (Espin et al., 2008)