A legibility scale for early primary handwriting: Authentic task and cognitive load influences

Cornelia Staats, Grace Oakley and Ida Marais
The University of Western Australia

This study set out to examine the range of legibility demonstrated by Western Australian students required to handwrite tasks of increasing intrinsic cognitive load. A representative sample of students in Years 1, 2 and 3 (N=437) was recruited for a cross sectional study and teachers administered handwriting tasks. Year 1 students were administered easier tasks (copying from the board and dictation), and Year 3 students were administered more difficult tasks (dictation and composition), whilst students in Year 2 were administered all three tasks. A rubric was then constructed for six aspects of legibility from selected participant exemplars: letter formation, size, space in word, space between words, line placement, and slant, providing 18 items for analysis (3 tasks x 6 aspects). The rubric demonstrated acceptable inter- and intra-reliability. Scores were assigned following pairwise comparisons; a Rasch model (RM) analysis was applied to scores. Fit to the RM was confirmed to permit a more accurate assessment of legibility. The study substantiates many assumptions about handwriting in the extant literature, and more specifically reveals how cognitive load governs legibility when students are learning to handwrite. Implications for practice are discussed.

Introduction

Handwriting in alphabetic systems is the graphic representation of written language by hand, but instruction is necessary to ensure students develop legible and fluent handwriting (Santangelo & Graham, 2016). Current research stresses the need for explicit, sequential teaching about where to place letters (e.g., ascender and descender letters), correct motor sequencing of letter strokes (where to start the letter and where to go next), memory retrieval practice of the letterforms, and student self evaluation of handwritten letter(s) (Fancher, Priestley-Hopkins & Jefferies, 2018; Shaw, 2011). It is estimated that young students require at least 20 hours of explicit instruction in order to learn to handwrite (Hoy, Egan & Feder, 2011; McCarroll & Fletcher, 2017). Even then, the quality of student handwriting deteriorates when they are asked to compose a text as opposed to copying text or scribing dictated text (Graham, Struck, Santoro & Berninger, 2006). The purpose of this study was to examine the range of legibility demonstrated by Western Australian Years 1 and 2 students when they were required to handwrite authentic classroom writing tasks of increasing intrinsic cognitive load, using the Rasch model of measurement for data analysis. Comparing individual student performance across writing tasks while they are learning to handwrite can better inform diagnosis and intervention to enhance handwriting proficiency.

Background and literature

Learning to handwrite is cognitively demanding (McCarney, Peters, Jackson, Thomas & Kirby, 2013; McCutchen, 2006). Students have to retrieve an orthographic representation (i.e. a letterform) from long term memory (McCloskey & Rapp, 2017), hold it in working memory, at the same time identify the motor program for the selected allograph (letter case), establish the size and placement of the letter, and finally handwrite it (Graham, Harris & Fink, 2000). Motor programs for individual letters are mental models that specify the number of basic motor units and their spatiotemporal relationship (Palmis et al., 2017). Without an ability to: distinguish alphabetic symbols as letters in words, construct mental models as orthographic representations to which phonemes can be applied, and assemble a motor program to execute the letters, students will not be able to independently retrieve letterforms to handwrite (McCloskey & Rapp, 2017). Transcription skills (handwriting and spelling) not only demand cognitive resources of beginner writers, they also interfere with cognitive resources available for text generation (Graham et al., 2000). In turn, the complexity of writing tasks, or intrinsic cognitive load, interferes with handwriting legibility for students learning to handwrite (Bourdin & Fayol, 2000; Olive, Favart, Beauvais & Beauvais, 2009). This inherent circularity potentially limits effective evaluation of legibility for students learning to handwrite.

Cognitive load theory (CLT) (Sweller, Ayres & Kalyuga, 2011) provides a framework to explicate the cognitive demands managed during authentic classroom writing tasks for students learning to handwrite. CLT is an information processing theory concerned with learning acquired through the dynamic relationship between instruction and cognitive architecture (Sweller, van Merrienboer & Paas, 1998). The theory assumes a cognitive architecture made up of long-term memory or organised schema, with varying degrees of automation, and a limited working memory capacity that includes partially independent components to deal with visual and auditory information (Baddeley, 2003). Three types of cognitive load at the specific level have featured in the literature and are considered to be additive (Gerjets, Scheiter & Cierniak, 2009; Sweller et al., 2011): intrinsic CL (task complexity); extraneous CL (instructional design); and germane CL (motivation and mental effort required from the learner). According to the theory, learning involves conscious processing of information and requires considerable effort; however, with deliberate practice acquired schema can be used automatically with minimal to no conscious effort (McCutchen, 2006). Understanding occurs when all necessary interacting elements can be processed in working memory and skilled performance comes from acquiring automated schema (Sweller et al., 2011). When cognitive load is too high, access to long term memory and the ability to add information to long term memory schema is compromised; on the other hand, learning increases when cognitive load is decreased (Sweller et al., 2011). Intrinsic cognitive load, in the form of a hierarchical order for writing task complexity, was considered for this study (hereafter cognitive load).

In literate societies, assumptions about students' cognitive architecture prior to handwriting instruction are not uncommon or unreasonable because most students learn to handwrite with instruction. For example, it is reasonable to assume that students have developed fine motor ability to hold a pencil, can copy basic shapes, and have knowledge of the alphabet. Grade 1 students' handwriting is considered readable, in context, if at least 70% of letters are legible and these students can be expected to become more proficient with time (Case-Smith, Holland, Lane & White, 2012; Feder, Majnemer, Bourbonnais, Blayney & Morin, 2007). Generally, however, legibility for students who are learning to handwrite is highly variable (Feder et al., 2007; Graham et al., 2006) and this can be a challenge for measurement (Graham & Weintraub, 1996). Previous handwriting studies using both global and analytic assessment methods to measure performance have consistently identified letter formations, letter size, space between words, space within words, slant and line placement as aspects of legibility (Graham & Weintraub, 1996; Simner & Eidlitz, 2000). These aspects explained 96% of the variance between good and poor handwriters in a study with typically developing Grade 1 and 2 students carrying out copying and composition tasks (Graham et al., 2006). Research findings suggest that examination of legibility aspects on assigned writing tasks of increasing cognitive load (copying from the board, hereafter copying; dictation; and composition) might present an efficient and effective evaluation strategy for students learning to handwrite.

Learning to handwrite is a precursor to developing handwriting fluency or handwriting speed (Graham et al., 2006), and handwriting speed is often considered when studying legibility. However, speed appears to be a different construct to legibility because slower handwriting speed is influenced by student attention, focus and organisation (Rosenblum, Aloni & Josman, 2010), spelling ability (Sumner, Connelly & Barnett, 2013), and content knowledge (Graham & Harris, 2016). The largest increase in handwriting speed is generally in Grade 3 after students have learnt to handwrite (Overvelde & Hulstijn, 2011; Palmis, Danna, Velay & Longcamp, 2017). In this study, neither handwriting speed nor spelling, although acknowledged as transcription skills, were examined.

This study builds on, and extends, previous research on handwriting and cognitive influence (Bourdin & Fayol, 2000; Olive et al., 2009). Based on CLT, automatic access to organised schemas in long term memory for orthographic representation(s) and motor program(s) can be inferred from legibility by incrementally increasing cognitive load (copying, dictation, composition), which is managed in working memory (McCarney et al., 2013; McCutchen, 2006). As a result, diagnosis and intervention in handwriting could be better informed by comparing legibility between students, and between task performances for individual students. In this study, a coherent series of hierarchical writing tasks that incrementally increased cognitive load was designed to examine the effect on students' legibility when learning to handwrite.

In contrast to previous studies examining legibility in beginner handwriters, the Rasch model of measurement (RM) was adopted for data analysis. The RM is especially suited for analysing hierarchically ordered tasks and does not assume normal distributions (Rasch, 1960). Instead, the difficulty of the items is compared to the ability level of the students completing the items. If the students administered the items are composed of a representative sample to complete such items, and provided the data conform to the RM, then the resultant measures are reproducible from one testing situation to another. In traditional measurement models, the total score when standardised is assumed to be at interval level, even when data are ordinal. In the RM, instead of remaining an ordinal scale the raw scores are transformed into interval scale by a log-odds transformation of the data (logarithmic transformation of the odds for correctly achieving the item) (Bond & Fox, 2001, p. 124). The logit is the log-odds unit and the unit of measurement for the legibility scale construction. Equal interval measures place item difficulty (legibility aspects according to writing task) and person abilities (legibility performance for each task) on the same scale so they can be directly compared to each other. At the same logit, students have a 0.5 probability of achieving that item; students with greater proficiency have a greater probably of achieving the item and students with lower proficiency have a lower probability of achieving the item. A priori, according to the RM and CLT, students with legible composition would demonstrate legible dictation and copying; and, conversely, students with illegible copying would not be expected to demonstrate legible dictation. If that were the case, then that would be an anomaly and warrant further investigation. To date no analysis of legibility when students are learning to handwrite using the RM has been documented in the literature, as far as we are aware.

The research questions were:

1. What are the effects of cognitive load on legibility when students are learning to handwrite?
2. How do aspects of legibility differentiate handwriting performance when students are learning to handwrite?

Method

Participants

The study was conducted at the end of school Term 1; there are four terms in Australian schools. Because the study used a cross sectional design, and to capture the full range of handwriting performance when students are learning to handwrite from the beginning of Year 1 to the end of Year 2, beginning Year 2 students acted as proxy for end of Year 1 and beginning Year 3 students acted as proxy for end of Year 2. In Australia, where this study took place, students turn 6 in Year 1, 7 in Year 2 and 8 in Year 3.

In order to obtain student responses across the required ability range for Years 1 and 2, a representative sample was recruited. Students were recruited from 11 schools, both government and private fee-paying schools, within a 10 kilometre radius of the Perth Central Business District. At the time of the study, which took place prior to the introduction of the Australian National Curriculum, private schools determined their own curriculum related to handwriting instruction in contrast to a prescribed curriculum in government schools; therefore, it was of interest if differences could be detected. Ethics approval to conduct the study was obtained by the University of Western Australia (RA/4/1/2599) and all information requirements were met. Informed consent was obtained from all individual participants included in the study. School principals were informed of the study's purpose and design before giving their written consent to recruit participants. Student participants gave verbal assent and parents completed written consent forms. Participants were de-identified prior to scoring legibility and recording data. Table 1 depicts the distribution of students used in the analyses according to person factors: gender, year level, and school type.

Table 1: Participant sample according to gender, year level, and school type

Funding type	Year 1		Year 2		Year 3		Total
	Male	Female	Male	Female	Male	Female	
Public	46	56	34	28	34	25	222
Private	38	39	35	39	27	36	215
Total	84	95	69	67	61	61	437

Research design

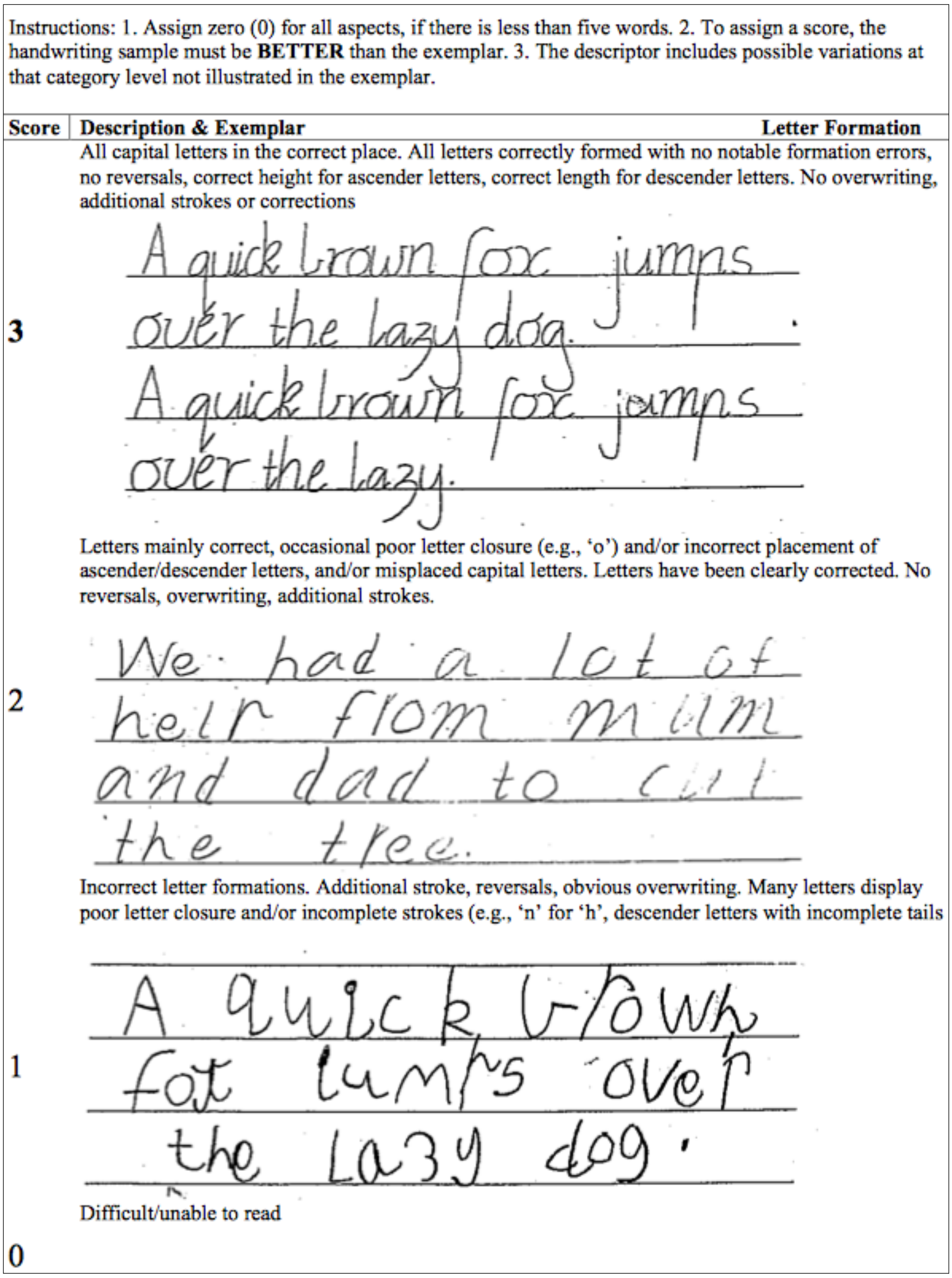
The study was conducted at the end of Term 1 as an assumption of the study was that all Year 1 students had received some handwriting instruction because the validity of results increases when students have received instruction (Borsboom, Mellenbergh & van Heerlen, 2004). At the time of the study, Year 1 was the first compulsory year of school in Western Australia. The intent of the study was not to discriminate between poor and good handwriters but, rather, to capture the range of legibility for hierarchically ordered authentic classroom writing tasks within students' ability range. A tailored design was adopted, a term used for selecting items based on their relative difficulties (Kline, 2015). Specifically, students in Year 3 were administered more difficult tasks (dictation and composition) and students in Year 1 were administered easier tasks (copying and dictation) with all three tasks administered to Year 2 providing a link among tasks. The assigned tasks accommodated the cross sectional design (copying and dictation at the beginning of Year 1, copying, dictation and composition at the end of Year 1, and dictation and composition at the end of Year 2). Table 2 illustrates the data structure for the tailored design of the legibility scale (LS).

Table 2: Data structure for tailored testing of the legibility scale

Year 1 (beginning Y1)	Year 2 (proxy end Y1)	Year 3 (proxy end Y2)
Copying	Copying	
Dictation	Dictation	Dictation
	Composition	Composition

Measures: Marking rubric

A marking rubric was constructed based on the following assumptions: first, based on CLT and an assumed cognitive architecture (Sweller et al., 2011), more legibility was evidence of more organised schemas for orthographic representations and motor programs (McCloskey & Rapp, 2017); second, each of the aspects would provide unique information to the construct of legibility and therefore performance variation in any one of the aspects would produce variation in measurement outcomes (Borsboom et al., 2004); and, finally, judgment by assessors was valid based on the law of comparative judgment which states that when comparing any two handwriting specimens, the magnitude of quality may not be measured directly, but can be inferred (Heldsinger & Humphry, 2010). Thorndike (1912) stated that the specimen being compared is either better or worse than the target specimen and therefore can be assigned a value; given specimen X, specimen Y is either X+Y or X-Y.



Instructions: 1. Assign zero (0) for all aspects, if there is less than five words. 2. To assign a score, the handwriting sample must be **BETTER** than the exemplar. 3. The descriptor includes possible variations at that category level not illustrated in the exemplar.

Score	**Description & Exemplar** **Letter Formation**
3	All capital letters in the correct place. All letters correctly formed with no notable formation errors, no reversals, correct height for ascender letters, correct length for descender letters. No overwriting, additional strokes or corrections A quick brown fox jumps over the lazy dog. A quick brown fox jumps over the lazy.
2	Letters mainly correct, occasional poor letter closure (e.g., 'o') and/or incorrect placement of ascender/descender letters, and/or misplaced capital letters. Letters have been clearly corrected. No reversals, overwriting, additional strokes. We had a lot of help from mum and dad to cut the tree.
1	Incorrect letter formations. Additional stroke, reversals, obvious overwriting. Many letters display poor letter closure and/or incomplete strokes (e.g., 'n' for 'h', descender letters with incomplete tails A quick brown fox jumps over the lazy dog.
0	Difficult/unable to read

Figure 1: LS rubric for letter formation (see Appendices 1 to 5 for LS rubrics for size, space in word, space between words, line placement and slant)

The rubric was constructed using participant samples. Exemplars were selected based on demonstrated difference in performance by means of paired comparison (Heldsinger & Humphry, 2010). The researcher rated pairs of work samples according to a predetermined schedule to choose successively better examples of the aspect under consideration. The rubric did not have structurally aligned categories such that each aspect had the same number of categories; instead, the number of exemplars to differentiate performance per aspect was not predetermined. For example, letter formation had four category levels but space within word had two category levels. Figure 1 illustrates the marking rubric for letter formation; other rubrics are given in Appendices 1 to 5. Selected handwriting exemplars acted as the threshold (category) for assigning a score; each exemplar was flanked by category descriptors (Humphry & Heldsinger, 2014). Based on six aspects of legibility and three writing tasks, there were 18 items to score. Inter-rater reliability, established by two raters (the researcher and an assistant) was satisfactory (Cohen's kappa=0.65), and intra-rater reliability (the researcher) was established a week later (Cohen's kappa=0.97). Internal consistency was good (alpha coefficient=0.88).

Data analysis

Data were analysed using the *Rasch Unidimensional Measurement Model* (RUMM2030) software (Andrich, Sheridan & Luo, 2012) according to the polytomous Rasch model (Andrich, 2009). Polytomous items are items that have more than two ordered categories. The RUMM program generates a number of statistics and graphical displays to decide whether or not the data fit the RM. Analogous to the traditional index of reliability, Cronbach coefficient alpha, the Rasch index of reliability, called the person separation index (PSI) has values between 0 and 1 with higher values indicating higher reliability (Andrich, 1982).

The model for dichotomous responses is illustrated by the item characteristic curve (ICC) and the model for polytomous items is illustrated by the category characteristic curve (CCC). Figure 2 shows the category characteristic curves for an item with 4 categories and a maximum score of 3. The CCC also illustrates points on the X-axis that are the thresholds (T1, T2, T3) between the successive categories. The threshold between two adjacent response categories is the point on the measurement continuum where the probability of a response in either of the two adjacent response categories is equal. Thresholds that are not ordered sequentially indicate that students had difficulty discriminating between them.

A 'family approach' to assessing item misfit was used (Smith & Plackner, 2009). An item was flagged as misfitting if it misfit according to two item fit statistics; the item fit-residual statistic and the item chi-square fit statistic. In addition, information from graphical displays as well as the content of the item were considered. Both the item fit-residual statistic and the item chi-square fit statistic are based on comparisons between observed and expected responses. Misfit is indicated if the probability value of the chi-square statistic is less than 0.01. The overall item/trait chi square statistic is the sum of the individual item chi square statistics and gives a summary indication of fit for the items as a group. In the case of the item residual fit statistic, misfit is generally indicated if the value is outside of the range -2.5 to 2.5. In addition, if the mean and SD of all item fit residuals are close to 0 and 1 respectively, good fit is indicated.

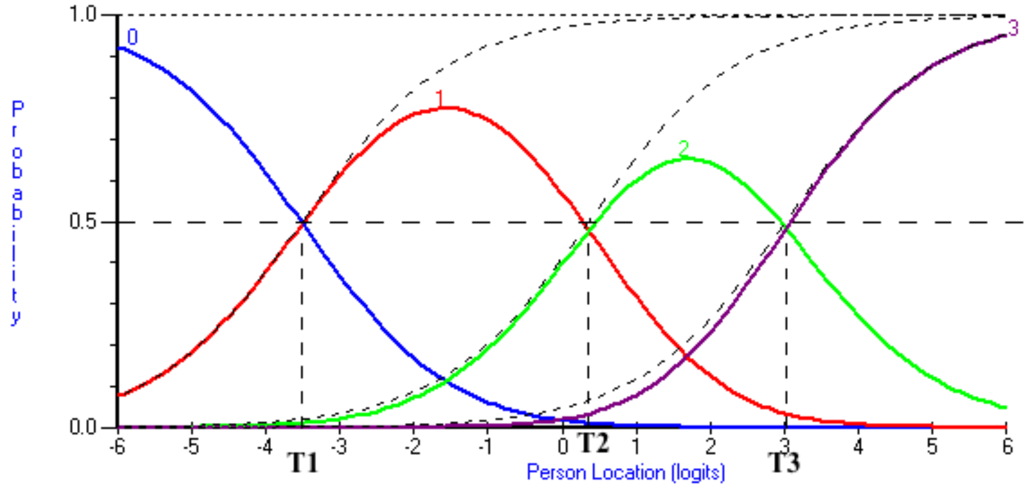


Figure 2: Category characteristic curve with four categories and three thresholds

Rasch models assume local independence of responses; that is, the response to one item does not depend on the response to another item. Two types of violations of local independence, response dependence and multidimensionality, were diagnosed by examining the item residual correlation matrix (Marais & Andrich, 2008). Item residual correlations between -0.3 to 0.3 are generally considered within accepted range.

Rasch analysis allows diagnosis of differential item functioning (DIF) for various sub-groups of the population. An item shows DIF when, for the same level of the trait being measured, members of one sub-group (e.g., males) score differently on the item than members of another sub-group (e.g., females). This does not preclude a different score between males and females on an item, but rather that, given the same overall level of handwriting legibility, the expected score on an item should be the same for different sub-groups. DIF is diagnosed graphically through an inspection of the item characteristic curve (ICC) and is confirmed statistically through an analysis of variance (ANOVA) of the residuals (Andrich, 2012). Items were tested for DIF for gender, handedness, type of school and year-level.

Procedure

Test prompts for the writing tasks were devised in consultation with experienced classroom teachers. The copying task was a pangram written on the whiteboard, *The quick brown fox jumped over the lazy dog.* The dictation task read, *We had a lot of help from Mum and Dad to cut the tree,* and the composition task prompt read, *My Family.* Time allowed for the copying and composition tasks was three and seven minutes respectively. Teachers administered tasks on consecutive days to avoid student fatigue in order of copying, dictation, and composition. No assistance was given during the writing tasks. Students were provided with standard 18 mm lined paper and used their own pencils. Year 1 students completed the copying and dictation tasks, Year 2 completed copying, dictation and composition tasks, and Year 3 completed dictation and composition tasks. All handwriting samples were scored by the researcher and scores entered into a spreadsheet. The item structure (3 tasks x 6 aspects=18 items) had propensity for a halo effect, so the order of scoring tasks was rotated for each classroom to minimise for this effect.

Students were removed from the data set for analysis if they were receiving additional support for handwriting, or had missing data for any of the writing tasks (as it is easier to study distribution and RM fit with complete data). Handwriting samples of less than five words were counted as missing data as some writing is necessary to assess handwriting proficiency. This left data for 437 students to be used in analyses (refer to Table 1 for participant factors).

Results

The RM for polytomous responses used data of 437 students to examine the overall fit of data to the RM in order to determine: (i) scalability (the validity of placing students and items on the same scale); (ii) hierarchal ordering of items and student ability (the relative difficulty of different legibility aspects and relative difficulty of different writing tasks; the relative ability levels of students); (iii) unidimensionality (summed scores confirm legibility construct); (iv) and item invariance (whether or not items functioned the same way the same for different sub-groups of the population).

Scale construction: Person-item distribution

The person-item distribution, alignment of items and persons, is depicted in Figure 3. The graph show histograms of the Rasch *person* estimates (top histograms for the different year-levels in blue, red and green) and *item* difficulty threshold estimates (bottom histogram in blue). The graph shows relatively good targeting of the test prompts with thresholds of difficulty across the whole continuum indicating that the test was generally within the ability range of the students. There were no disordered category thresholds, nor ceiling or floor effects. The separation between mean year levels (-1.56, 0.77, 1.73 logits) for Years 1, 2 and 3 respectively supports the validity of the marking rubric. The distribution of the year groups shows the mean value of year levels increases as expected, but there is a great deal of variation within each year level. Standard deviations were (1.35, 1.57. 1.46) for Years 1, 2, and 3 respectively. The overall range for legibility in Term 1 for Year 1 to Year 3 was -4.69 to 5.05 logits. The person separation index (PSI) was 0.90 and indicated excellent reliability and power to detect misfit.

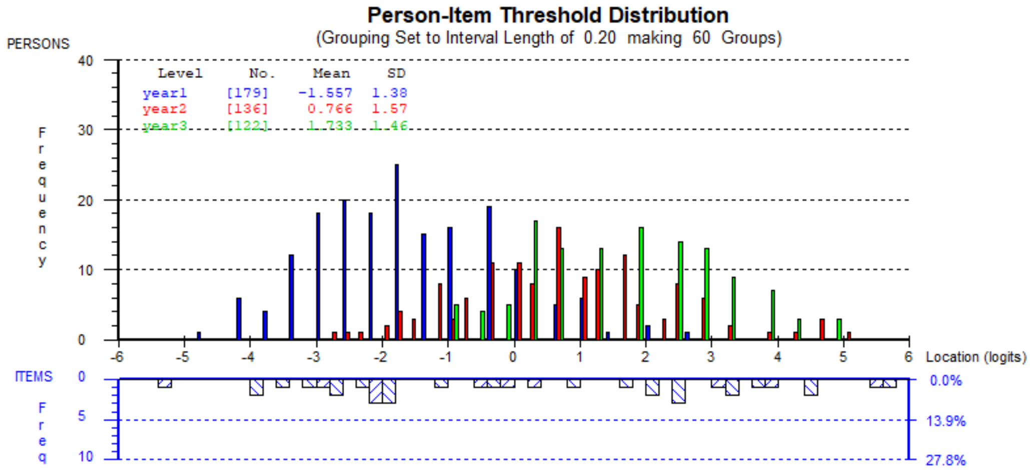


Figure 3: Person-item thresholds distribution according to year level
(use the 'zoom in' function for your PDF reader if necessary)

Fit of items and persons to the model

The overall item/trait chi square statistic value (□2=149.98, df 54; p<0.001), and the mean of the item fit residual statistics value for items (-0.57) and standard deviation (2.12) indicated there was some item misfit. The person residual fit statistic mean (-0.34) and standard deviation (1.01) were close to expected values of 0.00 and 1.00 indicating that the persons fitted the model. Item locations, standard errors and fit statistics are shown in Table 3. The table is ordered by the difficulty location of the items.

The following items met one criteria of misfit: Items 1, 2, 7, 9 and 13. Item 9 had a fairly high positive fit residual statistic (2.47) and showed poor discrimination between students (see Table 3). Figure 4 shows the ICC for item 9. As handwriting legibility increases along the X-axis, the observed means for students (black dots) did not increase as much as expected (curve).

Four items showed high negative fit residuals: Item 1 (z=-4.98), Item 2 (z=-2.92), Item 7 (z=-3.32), and Item 13 (z=-2.98). Items 1, 7 and 13 were for aspect letter formation and all showed very high discrimination, which suggests letter formation represents some higher-order feature of legibility. No item showed misfit according to both fit statistics and, after consideration of the graphic displays as well as the content of the items, all items were retained to maintain the structure of the test.

Hierarchical ordering: The relative difficulty of the tasks and items

Table 3 is ordered by the difficulty location of the items and the order of difficulty for writing tasks shows that increased cognitive load governs handwriting legibility. Copying

Table 3: Item locations of legibility aspects

Item no	Aspect	Location (logits)	SE	Fit residual	Chi square	p
9	Copy space in word	-2.71	0.17	2.47	44.68	0.00
3	Dictation space in word	-1.83	0.14	0.17	4.09	0.25
12	Copy slant	-1.81	0.16	-2.27	0.89	0.83
8	Copy size	-0.91	0.15	-0.97	2.75	0.43
6	Dictation slant	-0.73	0.12	-1.43	4.15	0.25
10	Copy space between words	-0.40	0.13	2.62	10.89	0.01
15	Compose space in word	-0.35	0.17	0.76	5.12	0.16
2	Dictation size	-0.27	0.11	-2.92	10.79	0.01
7	Copy letter formation	-0.04	0.11	-3.32	7.02	0.07
4	Dictation space between words	0.24	0.10	0.86	9.24	0.03
18	Composition slant	0.48	0.17	0.01	5.22	0.16
11	Copy line placement	0.60	0.13	1.04	4.02	0.26
14	Composition size	0.80	0.16	-0.04	1.68	0.64
5	Dictation line placement	1.13	0.10	2.04	11.69	0.01
1	Dictation letter formation	1.17	0.09	-4.98	12.25	0.01
16	Composition space between words	1.20	0.18	-1.03	3.07	0.38
13	Composition letter formation	1.45	0.12	-2.98	9.34	0.03
17	Composition line placement	1.97	0.14	-0.23	3.09	0.38

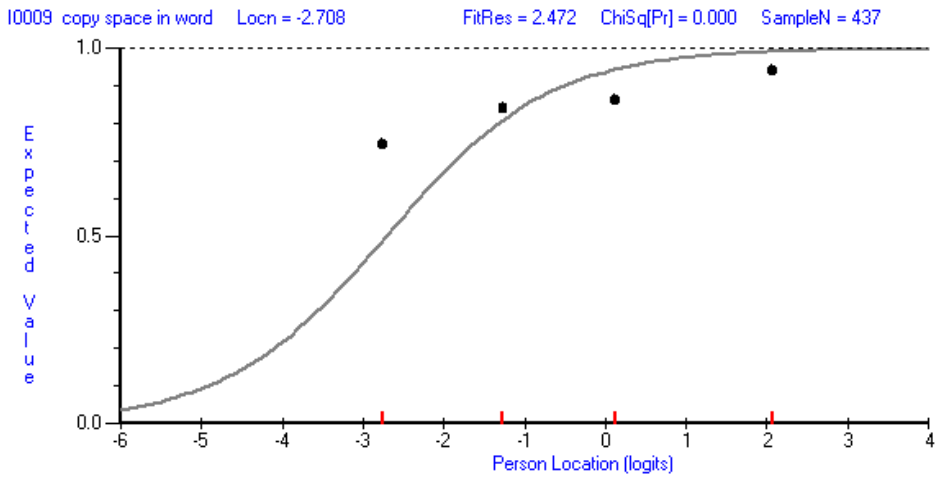


Figure 4: Item characteristic curve (ICC) for Item 9 - copying space in word

(M, -0.88) was easier than dictation (M, -0.04), which was easier than composition (M, 0.93). The specific mean logit values for copying, dictation and composition cannot be compared as ratios because there is no natural origin, but differences between values expressed in logits can be compared as ratios. The respective means of composition, dictation and copying were 0.93, -0.04, and -0.88. The successive differences between them were 0.97 and 0.84, giving a ratio of 0.97/0.84 = 1.15. There is a hierarchical proficiency for writing tasks but it is uneven. Dictation is more difficult than copying, but composition is much more difficult than dictation. The order for aspects from easy to difficult, expressed in logits according to mean locations was: space in word (-1.63); slant (-0.69); size (-0.23); space between words (0.37); letter formation (0.86); and line placement (1.25).

Uni-dimensionality and response dependence

The residual correlation matrix revealed some response dependence between letter formation, line placement, space in word, and slant, but it did not exceed 0.24; therefore, the magnitude of response dependence was considered relatively small. A unidimensional construct of legibility was upheld. Each item contributed uniquely to a valid measurement of legibility for students learning to handwrite in Years 1and 2.

Invariance: Differential item functioning (DIF)

No individual item showed DIF for the variables school type and year-level. One item showed DIF for gender. For the same level of legibility proficiency, girls were better at letter formation during composition (Item 13) than boys (ANOVA mean square=15.10). Based on this finding, composition must represent a slightly different writing task for boys than it does for girls, which is reflected in better letter formation. Forming letters legibly during composition appears more difficult for boys learning to handwrite. One item showed DIF for handedness: for the same level of legibility proficiency left-handed students did less well on Item 6 (dictation slant) than right-handed students (ANOVA Mean Square=11.65). Table 4 shows the differences between the mean person estimates for the different sub-groups.

Table 4: Subgroup mean differences, standard deviation, F-statistic and probability by person factor

Person factor		Number	Mean	Std dev.	F	p
Gender	Male	216	-0.23	1.99	10.15	0.001
	Female	221	0.39	2.04		
School type	Private	215	0.134	2.18	0.27	n.s.
	Government	222	0.034	1.88		
Handedness	Right	407	0.05	2.03	1.85	n.s.
	Left	30	0.57	2.03		
Year level	Year 1	179	-1.56	1.38	204.42	<0.001
	Year 2	136	0.77	1.57		
	Year 3	122	1.73	1.46		

Consistent with the literature that shows girls outperform boys in writing tasks during Grade 1 and 2, girls (M=0.39) did better than boys (M=-0.23) overall. As expected, the differences between year levels were statistically significant. There were no overall differences between government and private schools. Overall, there were no differences between left and right-handed students.

Summary of the Rasch model analysis

There was sufficient fit to the Rasch model to confirm the validity of the LS when students are learning to handwrite. Writing tasks with greater cognitive load had higher difficulty estimates, and those with less cognitive load were estimated to be easier, which supports previous research and the study's *a priori* assumption. Each of the six aspects contributed to a unidimensional scale of legibility, and the relative difficulty of each aspect was located on a linear scale. Test items operated in the same way for all participants with the caveat that composition letter formation operated differently for girls than for boys. The high order nature of letter formation when learning to handwrite was reinforced because it identified as a relatively difficult item that highly discriminated between year levels. In addition, a finding not reported elsewhere, is that line placement is a difficult legibility aspect for Year 1 and 2 students, even after letter formations have been acquired.

Discussion

The purpose of the study was to examine the range of legibility demonstrated by Years 1 and 2 students who were required to handwrite authentic writing tasks of increasing complexity (thereby increasing intrinsic cognitive load), using the Rasch model of measurement for data analysis. A specifically designed marking rubric was constructed based on selected participant exemplars, each showing increasingly better performance to define a threshold for assigning a score. Any one writing task was compared to selected exemplars for each one of the six aspects and assigned a total score for that task (task x aspect=item). Data obtained using the rubric shows sufficient fit to the Rasch model to describe variation between persons and items on a uni-dimensional scale of legibility, and demonstrates the utility of pairwise comparison to construct marking rubrics of performance (Heldsinger & Humphry, 2010).

The answer to the first research question, "What are the effects of cognitive load on legibility when students are learning to handwrite?", was illustrated by the establishment of a uni-dimensional scale for legibility. The effects of cognitive load on legibility for Year 1 and Year 2 students can now be examined in more detail. Although legibility differences according to writing task complexity have been reported elsewhere (Feder et al., 2007; Graham et al., 2006), the LS quantifies the magnitude of difference compared as ratios between copying, dictation and composition. The respective means of composition, dictation and copying were 0.93, -0.04, and -0.88. The successive differences between them were 0.97 and 0.84 giving a ratio of 0.97/0.84 = 1.15, confirming that each kind of writing task requires relevant instruction to the task when learning to handwrite (Graham et al., 2006). Dictation is more difficult than copying, but composition is much more difficult than dictation.

The LS did not distinguish between private and government schools, and as no data were collected on instructional methods, the findings suggest that they were effectively equivalent. No differences were detected between left- and right-handed students, confirming previous research (Graham et al., 2006). A significant difference was found between boys and girls, with girls being better than boys, a finding consistent with other studies (Berninger, Nielson, Abbott, Wijsman & Raskind, 2008; Malpique et al., 2017). Boys demonstrate more sensory motor immaturity than girls, including for fine motor ability (Larson et al., 2007) that may negatively impact learning to handwrite (Hooper et al., 2011). Other studies have suggested that boys are more likely to have problems with auditory processing, which negatively impacts processing speech sounds necessary for taking dictation (Rowe, Rowe, & Pollard, 2004).

The answer to the second research question, "How do aspects of legibility differentiate handwriting performance when students are learning to handwrite?", was addressed by examining the logit position of the six aspects of legibility on the LS. The order of difficulty for aspects of legibility from easy to hard was: space in word, slant, letter size, space between words, letter formation and line placement. Copying space in word and dictation space in word were very easy items and copying space in word discriminated poorly between persons. It might be that the aspect, space in word, is most applicable to composition as composition is a high cognitive load task and more likely to expose 'gaps' in cognitive architecture for writing. Slant was also an easy item, consistent with previous research (Graham et al., 2006). Letter formation strongly differentiated year levels and represents some higher order feature of legibility that operates differently at the beginning of Term 1 for students in Years 1, 2 and 3. At the same level of legibility proficiency, girls were better at letter formation during composition, which suggests composition is a slightly different task for girls than it is for boys, when both are learning to handwrite. The most difficult aspect was line placement. Whether or not to use lined or unlined paper depends on when it is introduced (Daly, Kelley & Krauss, 2003). When children start to experiment with writing letters, their attention is focused on reproducing letterforms instead of where to place them. On the other hand, lined paper provides a visual aid for maintaining horizontal consistency that may be differentially beneficial for poor handwriters, because they are more likely to demonstrate visual spatial difficulties (Fancher et al., 2018). Lined paper when students are receiving formal handwriting instruction is advised.

There were some limitations to the study. It was conducted with Year 1 students and most students now commence formal handwriting instruction one year earlier in Australia (school entry from 4.7 years), which may have implications for the findings (Malpique, Pino-Pasternak & Valcan, 2017). However, the LS showed no ceiling effects for Year 1 - 3 students in Term 1, which suggests that findings remain relevant despite this change in context. The study was cross sectional so the Year 2 study acted as a proxy for the end of Year 1 and Year 3 students acted as a proxy for the end of Year 2; therefore, the LS may not have captured the range as accurately as a longitudinal study. It is possible that the length of the summer break (7 weeks in Australia) could contribute to growth and change in Year 2 and Year 3 students that was not captured in a cross sectional study. The study was conducted in Western Australia and may not be immediately applicable to other jurisdictions. The rubric (Figure 1 and Appendices 1-5) could have included more exemplars (thresholds) to increase its precision. The study intentionally did not examine handwriting speed, as legibility is considered a precursor to developing handwriting speed.

Implications for practice

Four implications for practice emerge from the findings of the current study. First, the need for handwriting instruction to reduce variability is reinforced. The most critical time in the life of a reader or writer is a year after instruction has begun (Rowe et al., 2004, p. 30). Alphabet knowledge (letter recognition for upper and lower case, letter name, and sound-letter representations) contributes to mental orthographic representation of letter(s) that, in addition to its motor program, is necessary in order to handwrite (McCloskey & Rapp, 2017; Palmis et al., 2017). There is some evidence that at risk kindergarten children do not have well consolidated alphabet knowledge prior to handwriting instruction (Drouin & Harmon, 2009; Jones & Reutzel, 2012). Effective instructional practices that increase legibility integrate phonics instruction with handwriting (Shaw, 2011), and spelling instruction with handwriting (Graham, Harris & Adkins, 2018). Multisensory strategies may be useful at beginning stages of learning to handwrite (Shaw, 2011), but cognitive strategies (mental practice, self monitoring, self evaluation) are best at reducing letterform variations (Saleem & Gillen, 2019).

Second, the need to vary handwriting instruction according to the writing task is reinforced. Dictation is more difficult than copying, but composition is much more difficult than dictation; therefore, dictation can be considered a bridging task between copying and composition. Cognitive load can be manipulated to provide just the right challenge for students to maintain legibility when they are learning to handwrite. For example, when instructing students to take dictation students may be told to start the sentence with a capital letter. Instructors may say ‘space’ prior to dictating the following word and end the sentence with ‘full stop’. To reduce the cognitive load of spelling at the same time as maintaining legibility, instructors may ‘stretch’ out a word to help students hear each sound in order that they can ‘blend’ sounds as a word when handwriting. Instructional design for extended text may take the form of a five-sentence narrative or five-sentence information text that could be dictated one sentence at a time during daily handwriting lessons, to reinforce a salient feature(s) of legibility, and/or of text conventions (capital letter, full stop, space between words) at the same time. While composition as self generated text will always have a place at the beginning of the school year in Year 1 classrooms, it seems that using composition as a means of instruction and practice for handwriting legibility may be counterproductive.

Third, the use of pairwise comparison between writing tasks for the same student can be informative for diagnosis and intervention planning (Heldsinger & Humphry, 2010). The distinction between poor and good handwriters is often more obvious than for students ‘somewhere in the middle’ range, when students are learning to handwrite (Feder et al., 2007). Potential difficulties for individual students will become apparent as the cognitive load of writing tasks increases; the earlier this is detected, the earlier it can be addressed (Santangelo & Graham, 2016). The comparison of intra-individual responses on writing tasks of increasing cognitive load to assess handwriting proficiency in Years 1 and 2 is both efficient and effective and has been confirmed by this study.

Finally, monitoring students' handwriting while they are handwriting is encouraged (Barnett, Stainthorp, Henderson & Scheib, 2006). Early correction of directional or spatial placement mistakes can prevent unhelpful motor habits that are difficult to correct later. In addition, self evaluation by students of their own letterforms can improve legibility (Fancher et al., 2018; Vander Hart, Fitzpatrick & Cortesa, 2010), especially once students can independently retrieve the letter.

Conclusion

This study set out to examine the range of legibility that is demonstrated by students required to handwrite tasks of intrinsic cognitive load, using the Rasch model of measurement for data analysis. The findings confirm many assumptions held about handwriting instruction and handwriting legibility documented in the extant literature. The study adopted a novel approach to study legibility by devising writing tasks informed by cognitive load theory. The data were analysed by the Rasch model that placed students and items on a single scale and enabled description of variation between students and items, in contrast to classical test theory, which focuses on describing variation in the population. Strong evidence for legibility as a unidimensional construct was upheld and reinforces that legibility assessed by comparing authentic writing tasks of hierarchically ordered cognitive load is more informative to determine proficiency than single task evaluations when students are learning to handwrite. The comparison of intra-individual handwritten responses on writing tasks of increasing cognitive load aids diagnosis and intervention for handwriting legibility in Years 1 and 2.

References

Andrich, D. (1982). An index of person separation in latent trait theory, the traditional KR-20 index, and the Guttman scale response pattern. *Education Research and Perspectives,* 9(1), 95-104. https://www.rasch.org/erp7.htm

Andrich, D. (2009). Educational measurement: Rasch models. In E. Baker, B. McGaw & P. Peterson (Eds.) *International encyclopedia of education*, 3rd ed. Elsevier.

Andrich, D. & Hagquist, C. (2012). Real and artificial differential item functioning. *Journal of Educational and Behavioral Statistics,* 37(3), 387-416. https://doi.org/10.3102/1076998611411913

Andrich, D., Sheridan, B. & Luo, G. (2012). RUMM 2030. Perth, Australia: RUMM Laboratory. http://www.rummlab.com.au

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews/Neuroscience,* 4, 829-839. https://www.nature.com/articles/nrn1201

Barnett, A., Stainthorp, R., Henderson, S. E. & Scheib, B. (2006). *Handwriting policy and practice in English primary schools.* London: Institute of Education, University of London.

Berninger, V., Nielson, K. H., Abbott, R. D., Wijsman, E. & Raskind, W. (2008). Gender differences in severity of writing and reading difficulties. *Journal of School Psychology,* 46(2), 151-172. https://doi.org/10.1016/j.jsp.2007.02.007

Bond, T. G. & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences.* Mahwah, NJ: Lawrence Erlbaum Associates.

Borsboom, D., Mellenbergh, G. J. & van Heerden, J. (2004). The concept of validity. *Psychological Review,* 111(4), 1061-1071. https://doi.org/10.1037/0033-295X.111.4.1061

Bourdin, B. & Fayol, M. (2000). Is graphic activity cognitively costly? A developmental approach. *Reading & Writing,* 13(3-4), 183-196. https://doi.org/10.1023/A:1026458102685

Case-Smith, J., Holland, T., Lane, A. & White, S. (2012). Effect of a coteaching handwriting program for first graders: One-group pretest-posttest design. *The American Journal of Occupational Therapy*, 66(4), 396. https://doi.org/10.5014/ajot.2012.004333

Daly, C. J., Kelley, G. T. & Krauss, A. (2003). Relationship between visual-motor integration and handwriting skills of children in kindergarten: A modified replication study. *American Journal of Occupational Therapy,* 57(4), 459-462. https://ajot.aota.org/pdfaccess.ashx?url=/data/journals/ajot/930150/459.pdf

Drouin, M. & Harmon, J. (2009). Name writing and letter knowledge in preschoolers: Incongruities in skills and the usefulness of name writing as a developmental indicator. *Early Childhood Research Quarterly,* 24(3), 263-270. https://doi.org/10.1016/j.ecresq.2009.05.001

Fancher, L. A., Priestley-Hopkins, D. A. & Jeffries, L. M. (2018). Handwriting acquisition and intervention: A systematic review. *Journal of Occupational Therapy, Schools & Early Intervention*, 11(4), 454-473. https://doi.org/10.1080/19411243.2018.1534634

Feder, K. P., Majnemer, A., Bourbonnais, D., Blayney, M. & Morin, I. (2007). Handwriting performance on the ETCH-M of students in grade one regular education program. *Physical & Occupational Therapy in Pediatrics,* 27(2), 43-62. https://doi.org/10.1300/J006v27n02_04

Gerjets, P., Scheiter, K. & Cierniak, G. (2009). The scientific value of cognitive load theory: A research agenda based on the structuralist view of theories. *Educational Pscyhology Review,* 21(1), 43-54. https://doi.org/10.1007/s10648-008-9096-1

Graham, S. & Harris, K. R. (2016). A path to better writing: Evidence-based practices in the classroom. *The Reading Teacher*, 69(4), 359-365. https://doi.org/10.1002/trtr.1432

Graham, S., Harris, K. R. & Adkins, M. (2018). The impact of supplemental handwriting and spelling instruction with first grade students who do not acquire transcription skills as rapidly as peers: A randomized control trial. *Reading & Writing,* 31(6), 1273-294. https://doi.org/10.1007/s11145-018-9822-0

Graham, S., Harris, K. R. & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology,* 92(4), 620-633. https://psycnet.apa.org/doi/10.1037/0022-0663.92.4.620

Graham, S., Struck, M., Santoro, J. & Berninger, V. W. (2006). Dimensions of good and poor handwriting legibility in first and second graders: Motor programs, visual-spatial arrangement, and letter formation parameter setting. *Developmental Neuropsychology,* 29(1), 43-60. https://doi.org/10.1207/s15326942dn2901_4

Graham, S. & Weintraub, N. (1996). A review of handwriting research: Progress and prospects from 1980 to 1994. *Educational Psychology Review,* 8(1), 7-87. https://doi.org/10.1007/BF01761831

Heldsinger, S. & Humphry, S. (2010). Using the method of pairwise comparison to obtain reliable teacher assessments. *The Australian Educational Researcher,* 37(2), 1-19. https://doi.org/10.1007/BF03216919

Hooper, S. R., Costa, L.-J. C., McBee, M., Anderson, K. L., Yerby, D. C., Knuth, S. B. & Childress, A. (2011). Concurrent and longitudinal neuropsychological contributors to written language expression in first and second grade students. *Reading and Writing,* 24(2), 221-252. https://doi.org/10.1007/s11145-010-9263-x

Hoy, M. M. P., Egan, M. Y. & Feder, K. P. (2011). A systematic review of interventions to improve handwriting. *Canadian Journal of Occupational Therapy,* 78(1), 13-25. https://doi.org/10.2182/cjot.2011.78.1.3

Humphry, S. M. & Heldsinger, S. A. (2014). Common structural design features of rubrics may represent a threat to validity. *Educational Researcher,* 43(5), 253-263. https://doi.org/10.3102/0013189x14542154

Jones, C. D. & Reutzel, D. R. (2012). Enhanced alphabet knowledge instruction: Exploring a change of frequency, focus, and distributed cycles of review. *Reading Psychology,* 33(5), 448-464. https://doi.org/10.1080/02702711.2010.545260

Kline, P. (2015). *A handbook of test construction (psychology revivals): Introduction to psychometric design*. London Routledge.

Larson, J. C. G., Mostofsky, S. H., Goldberg, M. C., Cutting, L. E., Denckla, M. B. & Mahone, E. M. (2007). Effects of gender and ages on motor exam in typically developing children. *Developmental Neuropsychology,* 32(1), 543-562. https://doi.org/10.1080/87565640701361013

Malpique, A. A., Pino-Pasternak, D. & Valcan, D. (2017). Handwriting automaticity and writing instruction in Australian kindergarten: An exploratory study. *Reading & Writing,* 30(8), 1789-1812. https://doi.org/10.1007/s11145-017-9753-1

Marais, I. & Andrich, D. (2008). Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *Journal of Applied Measurement,* 9(3), 200-215. https://www.ncbi.nlm.nih.gov/pubmed/18753691

McCarney, D., Peters, L., Jackson, S., Thomas, M. & Kirby, A. (2013). Does poor handwriting conceal literacy potential in primary school children? *International Journal of Disability, Development and Education*, 60(2), 105-118. https://doi.org/10.1080/1034912X.2013.786561

McCarroll, H. & Fletcher, T. (2017). Does handwriting instruction have a place in the instructional day? The relationship between handwriting quality and academic success. *Cogent Education*, 4(1), 1-11. https://doi.org/10.1080/2331186X.2017.1386427

McCloskey, M. & Rapp, B. (2017). Developmental dysgraphia: An overview and framework for research. *Cognitive Neuropsychology,* 34(3/4), 65-82. https://doi.org/10.1080/02643294.2017.1369016

McCutchen, D. (2006). Cognitive factors in the development of children's writing. In C. MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 115-130). New York: Guilford Press.

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. https://doi.org/10.1016/j.learninstruc.2008.05.005

Overvelde, A. & Hulstijn, W. (2011). Handwriting development in grade 2 and grade 3 primary school children with normal, at risk, or dysgraphic characteristics. *Research in Developmental Disabilities,* 32(2), 540-548. https://doi.org/10.1016/j.ridd.2010.12.027

Palmis, S., Danna, J., Velay, J.-L. & Longcamp, M. (2017). Motor control of handwriting in the developing brain: A review. *Cognitive Neuropsychology,* 34(3/4), 187-204. https://doi.org/10.1080/02643294.2017.1367654

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research

Rosenblum, S., Aloni, T. & Josman, N. (2010). Relationships between handwriting performance and organizational abilties among children with and without dysgraphia: A preliminary study. *Research in Developmental Disabilities,* 31(2), 502-509. https://doi.org/10.1016/j.ridd.2009.10.016

Rowe, K., Rowe, K. & Pollard, J. (2004). 'Literacy behaviour' and auditory processing: Building 'fences' at the top of the 'cliff' in preference to ambulance services at the bottom. In *Supporting student wellbeing: What does the research tell us about social and emotional development of young people?* ACER Research Conferences. http://research.acer.edu.au/research_conference_2004/6

Saleem, G. T. & Gillen, G. (2019). Mental practice combined with repetitive task practice to rehabilitate handwriting in children. *Canadian Journal of Occupational Therapy*, (online first). https://doi.org/10.1177/0008417418824871

Santangelo, T. & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review,* 28(2), 225-265. https://doi.org/10.1007/s10648-015-9335-1

Shaw, D. M. (2011). The effect of two handwriting approaches: D'Nealian and Sunform, on kindergartners' letter formations. *Early Childhood Education Journal,* 39(2), 125-132. https://doi.org/10.1007/s10643-011-0444-2

Simner, M. L. & Eidlitz, M. R. (2000). Work in progress: Toward an empirical definition of developmental dysgraphia: Premliminary findings. *Canadian Journal of School Psychology,* 16(1), 103-110. https://doi.org/10.1177/082957350001600108

Smith, R. M. & Plackner, C. (2009). The family approach to assessing fit in Rasch measurement. *Journal of Applied Measurement,* 10(4), 424-437. http://jampress.org/abst2009.htm

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

Sweller, J., Ayres, P. & Kalyuga, S. (2011). *Cognitive load theory.* New York: Springer.

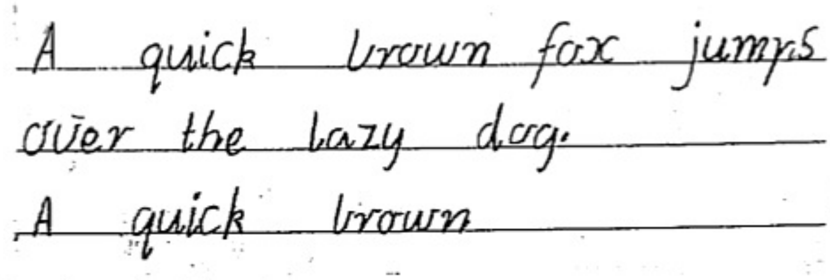
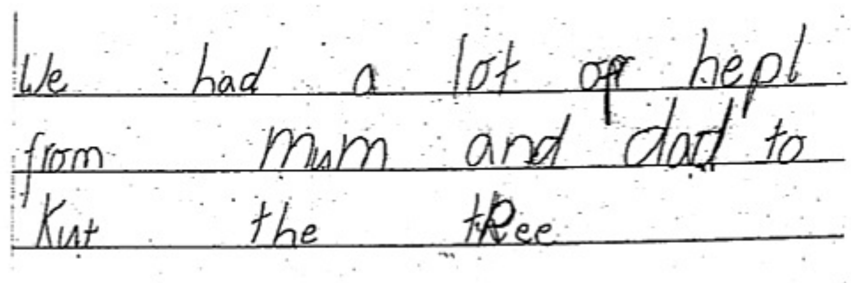
Sweller, J., van Merrienboer, J. J. & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review,* 10(3), 251-296. https://doi.org/10.1023/A:1022193728205

Thorndike, E. L. (1912). *Handwriting.* Teachers College, Columbia University, New York. https://archive.org/details/handwriting01thor/page/n6

Vander Hart, N., Fitzpatrick, P. & Cortesa, C. (2010). In-depth analysis of handwriting curriculum and instruction in four kindergarten classrooms. *Reading and Writing,* 23(6), 673-699. https://doi.org/10.1007/s11145-009-9178-6

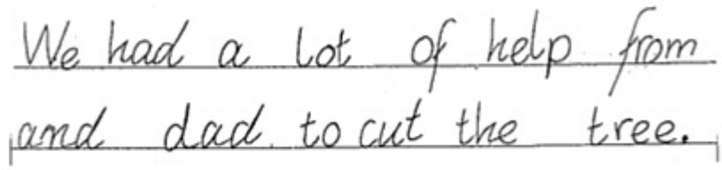
Appendix 1: LS Rubric: Size

Instructions: 1. Assign zero (0) for all aspects, if there is less than five words. 2. To assign a score, the handwriting sample must be **BETTER** than the exemplar. 3. The descriptor includes possible variations at that category level not illustrated in the exemplar.

Score	Description & Exemplar	Size
2	Letter size mainly consistent 	
1	Letter size moderately inconsistent 	
0	Letter size highly inconsistent	

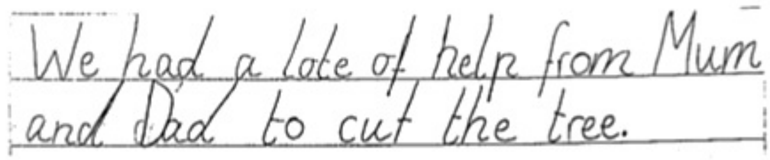
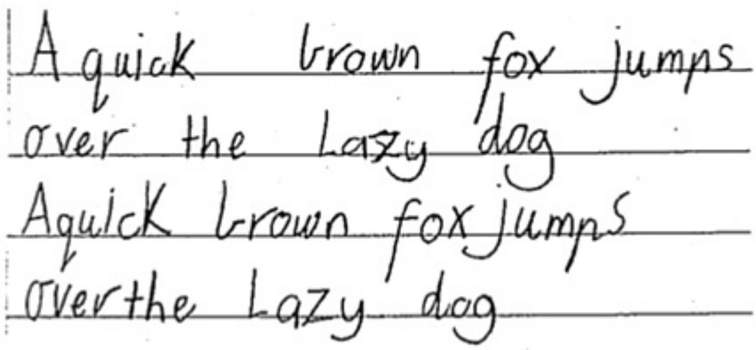
Appendix 2: LS Rubric: Space in word

Instructions: 1. Assign zero (0) for all aspects, if there is less than five words. 2. To assign a score, the handwriting sample must be **BETTER** than the exemplar. 3. The descriptor includes possible variations at that category level not illustrated in the exemplar.

<table>
<tr><th>Score</th><th>Description & Exemplar</th><th>Space in Word</th></tr>
<tr><td>1</td><td>Within words, no letters touch or overlap. Cursive exit and entry strokes may touch but NOT overlap
</td><td></td></tr>
<tr><td>0</td><td>Within words, letters touch or overlap</td><td></td></tr>
</table>

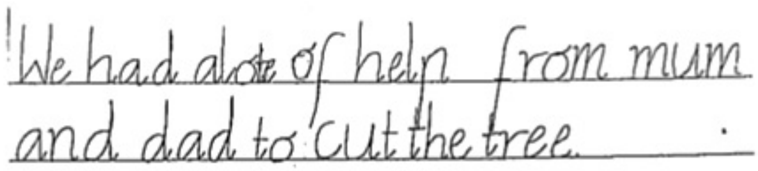
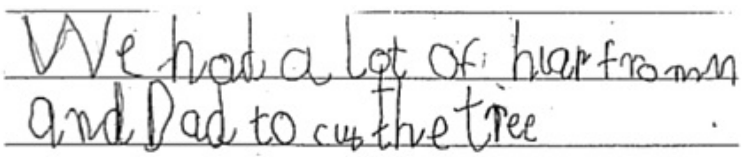
Appendix 3: LS Rubric: Space between words

Instructions: 1. Assign zero (0) for all aspects, if there is less than five words. 2. To assign a score, the handwriting sample must be **BETTER** than the exemplar. 3. The descriptor includes possible variations at that category level not illustrated in the exemplar.

Score	Description & Exemplar	Space between Words
2	Mainly consistent spaces between distinguishable words 	
1	Obviously inconsistent spaces between distinguishable words 	
0	Poor or distinct spaces between words	

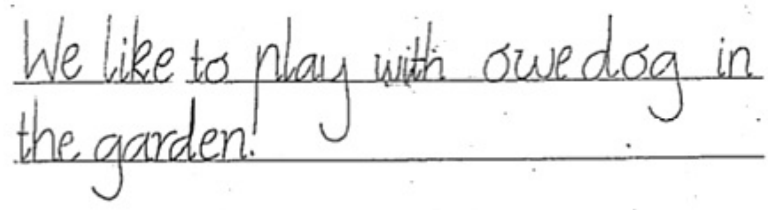
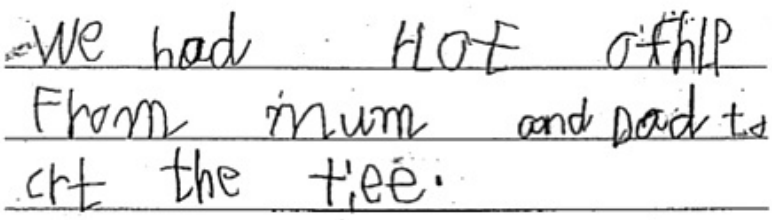
Appendix 4: LS Rubric: Line placement

Instructions: 1. Assign zero (0) for all aspects, if there is less than five words. 2. To assign a score, the handwriting sample must be **BETTER** than the exemplar. 3. The descriptor includes possible variations at that category level not illustrated in the exemplar.

Score	Description & Exemplar Line Placement
2	Consistent line placement. **Letters** are on the line 
1	Inconsistent line placement. **Letters** are either under or above the line. 
0	Words clearly 'trail'off/onto the line. Words are written between the lines

Appendix 5: LS Rubric: Slant

Instructions: 1. Assign zero (0) for all aspects, if there is less than five words. 2. To assign a score, the handwriting sample must be **BETTER** than the exemplar. 3. The descriptor includes possible variations at that category level not illustrated in the exemplar.

<table>
<tr><th>Score</th><th>Description & Exemplar Slant</th></tr>
<tr><td>2</td><td>Mainly regular slant
</td></tr>
<tr><td>1</td><td>Moderately irregular slant
</td></tr>
<tr><td>0</td><td>Highly irregular slant</td></tr>
</table>

Dr Cornelia Staats (corresponding author) is a paediatric occupational therapist in private practice. She works with children with developmental vulnerabilities including learning difficulty. She completed her doctoral studies in the Graduate School of Education at The University of Western Australia.
Email: cstaats@iinet.net.au

Dr Grace Oakley is Associate Professor in the Graduate School of Education at The University of Western Australia where she teaches, supervises and researches in the substantive area of literacy and educational technology.
Email: grace.oakley@uwa.edu.au
Web: https://research-repository.uwa.edu.au/en/persons/grace-oakley

Dr Ida Marais is a Senior Research Fellow in the Graduate School of Education at The University of Western Australia where she teaches the Rasch measurement of modern test theory. Her research area is the use of the Rasch model as psychometric method for analysis of test and questionnaire data.
Email: ida.marais@uwa.edu.au
Web: https://research-repository.uwa.edu.au/en/persons/ida-marais

Please cite as: Staats, C., Oakley, G. & Marais, I. (2019). A legibility scale for early primary handwriting: Authentic task and cognitive load influences. *Issues in Educational Research*, 29(2), 537-561. http://www.iier.org.au/iier29/staats.pdf