Regular Article

Journal of Psychoeducational Assessment
2026, Vol. 0(0) 1–17

DOI: 10.1177/07342829261416020
journals.sagepub.com/home/jpa
Sage

Detailed Assessment of the Speed of Handwriting – 2nd Edition (DASH-2): Factor Structure and Measurement Invariance Across Age and Sex

Anna L. Barnett[1], Sheila E. Henderson[2], Beverly Scheib[3], and Louis-Charles Vannier[4]

Abstract

The Detailed Assessment of the Speed of Handwriting-2 combines and extends previous versions of the test. This study examined the factor structure and presents the first analysis of measurement invariance across two age groups (8–16, 17–25) and sex. The normative sample comprised 762 examinees from the UK, Australia, and New Zealand who completed five tasks: Copy Best (CB), Alphabet Writing (AW), Copy Fast (CF), Free Writing (FW), and Graphic Speed (GS). As there were no significant between-country differences, confirmatory factor analysis (CFA) was performed on the combined data set. This supported the two-factor model corresponding to the test structure of four core tasks (CB, AW, CB, and FW) reflecting handwriting ability, with a separate measure of grapho-motor ability (GS). Multigroup CFA indicated full configural, metric, and scalar invariance for age groups and sex. The findings provide support for the test structure and reliability of the test across age and sex.

Keywords

writing assessment, dysgraphia, dyslexia, developmental coordination disorder (DCD), copying, DASH17+

---

[1] Centre for Psychological Research, Oxford Brookes University, UK
[2] Retired UCL Institute of Education, Hartford, Huntingdon, CAMBS, UK
[3] Independent Scholar/Researcher, Stoke D'Abernon, Cobham, Surrey, UK
[4] Pearson Clinical, Immeuble WeWork, Paris, France

Corresponding Author:
Anna L. Barnett, Centre for Psychological Research, Oxford Brookes University, Headington Campus, Oxford OX3 0BP, UK.
Email: abarnett@brookes.ac.uk

Introduction

Handwriting is recognized as a critical component of literacy and forms part of early school curricula across the world (e.g. in the UK (DfE, 2013) and Australia (ACARA, 2022)). Like reading and spelling, handwriting is a skill that needs to be taught, and these three components of literacy are interrelated. In order to develop their skill in writing, students must draw on their reading and spelling skills and teaching reading has been shown to enhance writing performance (Graham et al., 2018). Conversely, teaching writing has been shown to have a positive effect on competence in reading (Graham & Hebert, 2011), further emphasizing the reciprocal nature of the two skills. In particular, learning to write letters and words by hand has been found to benefit learning to spell and read (e.g. Pritchard et al., 2021; Wiley & Rapp, 2021). According to evidence from a recent meta-analysis by Araújo et al. (2022), one of the mechanisms underlying this link is that letter knowledge is enhanced by observation of the visual output of the motor act of handwriting. As students progress through education, writing demands increase and handwriting is used in a variety of ways, including to demonstrate knowledge, express ideas, and to take notes. Although other writing modes are used in classrooms, and typing in particular has become widely adopted, most class tests and the majority of formal high stakes examinations are still handwritten (Santangelo & Graham, 2016; Coombe et al., 2020). At this point in time, therefore, efficient handwriting continues to support success both within the educational system and in workplace settings that still require the ability to write legibly and fast (Bazerman et al., 2017).

Handwriting is part of the broader skill of writing, drawing on cognitive, language, perceptual, and motor skills which must be orchestrated to transform an idea into a written trace on the page. In some models of writing (e.g. Berninger & Winn, 2006), handwriting is simply viewed as a lower level ‘transcription’ skill (alongside spelling) with little detail on the processes involved. In the more focused psychomotor models of handwriting, however, Van Galen (1991) began by elaborating upon the transcription process by specifying the perceptual and motor processes involved in planning the shape and size of letter strokes. Since then, this model has been updated to acknowledge the influence of different language units (e.g. syllables) on aspects of handwriting production (Ahmed et al., 2022; Kandel et al., 2011).

The skill of learning to form letters and produce them consistently takes time. It has been suggested that achieving the point when the actions involved become ‘automatic’ in the sense that they require little focussed attention does not begin until the age of 8 years (Thibon et al., 2018). Given that different countries start to teach handwriting at different ages, however, using chronological age as a metric without taking account of years of experience might be difficult to interpret in this case. Nevertheless, regardless of the metric used, once the basic skill of forming letters is in place there are two aspects of handwriting, legibility and speed (or ‘fluency’), which continue to develop and show different developmental trajectories through childhood (Barnett et al., 2010; Gosse et al., 2021) with speed of production often being taken as a proxy for handwriting automaticity. As speed/automaticity increases, it is argued that the demands on executive functions (including working memory) are reduced and may be allocated instead to ‘higher level’ aspects of text production, resulting in higher quality text composition (e.g. Alves et al., 2016; Salas & Silvente, 2020). Consistent with these views is the idea that handwriting difficulties impede automaticity and affect higher order aspects of writing.

As handwriting relies on a variety of skills, individuals may experience difficulty for different reasons. Consequently, it is well documented that handwriting difficulties are common in children and young people with various neurodevelopmental disorders. This includes those with dyslexia (Hebert et al., 2018), Developmental Coordination Disorder (DCD; Barnett & Prunty, 2021), Attention Deficit/Hyperactivity Disorder (ADHD; Graham et al., 2016), and dysgraphia (Breaux & Munsell, 2025a; McCloskey & Rapp, 2018). Although the last term, ‘dysgraphia’, is not used in a uniform way in the literature (Chung et al., 2020), it is often associated with the DSM-5 classification of Specific Learning Disorder (with impairment in written expression, F81.81, American Psychiatric Association, 2022). In all of these groups, difficulties with legibility and speed are commonly reported but more specific differences in pen kinematics and pressure compared to peers have also been found (Gargot et al., 2020). For example, in children with dyslexia, poor spelling has been linked to slower handwriting, with excessive pauses when writing words (Sumner et al., 2014). In studies of children with DCD, Prunty and colleagues have found less accurate and less consistent letter formation (Prunty & Barnett, 2020) and more pausing during illegible words (Prunty et al., 2014) compared to age-matched peers. Brossard-Racine et al. (2015) examined handwriting difficulties in children with newly diagnosed ADHD. Although there were overall improvements with the use of stimulant medication, legibility remained persistently poor for over half of the sample.

Although there are individual differences in the nature and extent of handwriting difficulties both within and between these disorders, all can lead to poor outcomes in writing, broader literacy skills and beyond. As handwriting continues to be required for engagement in many classroom activities at all ages, difficulties in this area have been shown to have an impact on academic progress generally. For instance, faster, more automated handwriting has been linked to longer texts with greater compositional quality (Feng et al., 2019). This effect has been demonstrated even in primary-school aged children, where poor fluency was shown to impact statutory school assessment results (Medwell et al., 2009). Thus, students who can write more in a given time have the opportunity to demonstrate fuller knowledge and understanding. Assignments that are neater and easier to read also tend to be given higher marks (Graham et al., 2011). Children can be aware that they struggle to keep up with writing demands and to write legibly and this has been linked to lower perceived self-efficacy (Engel-Yeger et al., 2009). Psychological and emotional impacts of handwriting difficulties have also been found in children and adolescents (Hen-Herbst & Rosenblum, 2022; Zwicker et al., 2017).

Given the continued importance of handwriting through education and beyond, robust measurement tools are required to identify those with handwriting difficulties and plan appropriate support for them. One of the most popular tools for assessing writing in the English language is the Detailed Assessment of Speed of Handwriting (DASH; Barnett et al., 2007; Barnett et al., 2009), first developed to assess handwriting speed in 9- to 16-year-olds, then later extended to include 17- to 25-year-olds (DASH17+; Barnett et al., 2010; Barnett et al., 2011). Unlike other handwriting speed tests that include only one writing task (often copying, e.g., Killeen et al., 2006), DASH/DASH17+ contain a set of five tasks (see Table 1). This not only increases reliability of the total score but also gives an indication of performance under different writing demands, which in turn helps examiners understand how the examinee manages the various components of writing. For example, the 'Free Writing' task in DASH/DASH17+, which involves writing on a set topic (with topic prompts provided) for 10 minutes, places the greatest demands on the writer because it includes all elements of writing from idea generation, structuring a sentence and spelling words to controlling the pen to form legible letters on the page. Two sentence copying tasks (one in 'best' and one in 'fast' handwriting, each for 2 minutes) have lower demands because they provide the text, which must then be visually monitored and the appropriate letter forms and words generated. Comparison of the 'best' and 'fast' condition shows the extent to which the examinee can speed up and/or maintain legibility. Alphabet Writing (writing out the alphabet for 1 minute) involves retrieving the orthographic letter forms from memory in the correct sequence. Each of these tasks, defined as 'core' tasks, yields a scaled score which can then be combined to produce a total standardized score. The DASH/DASH17+ also include a task labelled as 'supplementary', which, unlike the other four tasks, does not involve a language component, as no letters or words are produced. This 1 minute 'Graphic Speed' task is designed to focus on the perceptual-motor aspects

Table 1. Description of the DASH-2 tasks and scores

	Stimulus material	Instructions	Measure	Mean scaled/ Standard score (SD)
Copy Best (CB) 2 minutes	Sentence containing all letters of the alphabet	Copy the sentence in 'best' handwriting	Mean wpm – number of legible words	10 (3)
Alphabet Writing (AW) 1 minute	None	Write the alphabet from memory in the correct sequence	Letters per minute – number of legible letters in correct sequence	10 (3)
Copy Fast (CF) 2 minutes	Sentence containing all letters of the alphabet (same as for Copy Best)	Copy the sentence quickly, making sure the words are readable	Mean wpm – number of legible words	10 (3)
Free Writing (FW) 10 minutes	Topic title with prompts presented as a 'spider diagram'	Write for 10 minutes	Mean wpm – number of legible words	10 (3)
Graphic Speed (GS) 1 minute (supplementary task)	Graphic Speed sheet containing circles	Quickly draw crossed lines in each circle, trying not to break the rules for accuracy	Number of correct crossed lines (X) produced in 1 minute	10 (3)

Note. Scores from the four 'core' tasks (CB, AW, CF, and FW) are used to produce a Total Standard Score (Mean = 100, SD = 15). The GS task provides a supplementary measure.

of handwriting and provides a measure of the ability to work quickly to control a pen to make two crossed lines within a given boundary. A scaled score for this task is also available.

As part of the initial development of a scoring system for DASH, correlations between all five DASH-2 tasks were first conducted using the standardization data (as reported in the test manual, Barnett et al., 2007). While the Graphic Speed task had rather low correlations with the other tasks, the other four tasks were substantially intercorrelated. A principal component analysis was then conducted to investigate the factor structure of the test at each year (9–16 years). In each case, only one strong component emerged, explaining between 57 and 69% of the total variance. Once again, Graphic Speed stood out from the other four tasks, with factor loadings ranging from 0.32 to 0.75 with five under 0.60. Factor loadings on the other four tasks were high across all age groups (ranging from .74 to .94). These findings suggested that four of the tasks could be regarded as a closely related homogenous set which could be combined to produce a meaningful total score. The findings also led to the decision to present the Graphic Speed task as a separate supplementary measure of motor control. The original publication of DASH was therefore based on this test structure.

When the test was extended for an older age group, examination of the tasks was repeated using the standardization data for DASH17+ (as reported in the test manual, Barnett et al., 2010). Similar findings emerged supporting use of the same test structure, with four 'core' tasks and Graphic Speed as a separate measure. Normative data for the DASH and DASH17+ were gathered in the UK. For both editions of the test, significant developmental trends were present with only minor differences between tasks. However, beyond the age of 17, rates of change slowed, such that norms for broader age groups seemed more appropriate: 17–18, 19–21, and 22–25 years.

A recently revised and updated version of the test, the Detailed Assessment of the Speed of Handwriting – 2nd Edition (DASH-2; Barnett et al., 2024), combines the two previous versions and provides normative data on an extended age range from 8 to 25 years. Recommendations from an expert panel and feedback from DASH/DASH17+ users from different professions contributed to the test revisions. These include a downward extension from 9 to 8 years, as users had found it suitable for use with younger children. New features include new sentences, designed for younger (8–16 years) and older (17–25 years) examinees, and alternative sentences for when repeat testing is needed. An alternative topic, created to provide a more challenging option for the Free Writing task was also provided. Norms were gathered between August 2021 and February 2023 in the UK, Australia, and New Zealand. Although traditionally single-country norms have been gathered, it is now more common for pan-country norms to be produced where the countries have similar demographics, education systems, and provision for students with special/additional needs, as was the case for the three regions included in the DASH-2 standardization project.

The purpose of a test manual is to provide the user with the reliability and validity data required to have confidence in the psychometric properties of the assessment tool they have chosen to use. Providing further evidence of these properties is then an ongoing process undertaken by researchers and test users should continue to consider the accumulated literature (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). For example, the new DASH-2 manual provides information on test-retest reliability for the two age groups completing the new sets of sentences, a younger group (aged 8–16 years) and an older group (aged 17–25 years), reporting good to excellent stability over time. The standard error of measurement is also provided for each of the tasks, as well as the Total Standard Score. Adding to evidence of test validity, the DASH-2 manual also reports data from special group studies which show the clinical sensitivity of the DASH-2 to the presence of difficulties with handwriting speed in a group with dyslexia and in a group with developmental coordination disorder (DCD). In terms of the test structure, reported intercorrelations between the five component tasks reflect findings from the previous versions of the test as described above, as do the results of a confirmatory factor analysis, which once again support the structure of the test in terms of the four 'core' tasks representing a multi-faceted concept of handwriting speed and a separate task measuring perceptual-motor control of a writing implement. However, given the wide age range for DASH-2, along with the introduction of some new materials, it is important to inform users of the validity of the test structure for both younger and older examinees. Furthermore, separate analyses for males and females would also be informative for users. The aim of the present study was therefore to examine further the factor structure as presented in DASH-2 and measurement invariance across a younger and older age group and across male and female participants.

Methods

Participants

The individuals included in this study comprised the normative sample employed in the standardization of the DASH-2. A stratified sample of 762 participants (377 males, 385 females) aged 8:00–25:11 years took part (see Table 2). For the project as a whole, a sampling plan was developed to ensure that representative proportions of participants from each country and demographic group in the UK, Australia, and New Zealand would be included. Data gathered from the 2011 Census in the UK, the 2021 Census in Australia, and the 2018 Census in New Zealand provided the basis for stratification. A total of 394 participants were selected from Australia/ New Zealand and 368 from the UK.

Within each country, the sampling plan involved a cell structure that identified appropriate numbers of participants for each cell, defined in terms of age, gender, geographic region,

Table 2. Distribution of sample by age and sex

	Male	Female	Total
	n	*n*	*n*
Younger age group			
8:0–8:11	50	52	102
9:0–9:11	51	49	100
10:0–11:11	53	62	115
12:0–14:11	63	52	115
15:0–16:11	50	51	101
Older age group			
17:0–18:11	34	37	71
19:0–21:11	35	38	73
22:0–25:11	41	44	85
Total	377	385	762

urbanicity, race/ethnic group/ancestry, and parental education. For the variable age, year groups were collapsed on the basis of analyses undertaken for previous editions of the test (see DASH-2 Administration and Scoring Manual, Barnett et al., 2024 for full demographic details).

Procedure

Ethical approval for the study was obtained from Pearson Clinical Scientific Council, Texas USA. Individuals aged 15 years or younger in the UK, and aged 17 years and younger in Australia and New Zealand, were registered for the project by a parent/carer, who gave informed consent. In Australia and New Zealand, young people aged 12–16 years were also given the option to provide their own informed consent. Those aged 16–25 years in the UK and 18–25 in Australia and New Zealand were able to provide their own informed consent.

Participants were individually tested on the five DASH-2 tasks administered in the following order: Copy Best, Alphabet Writing, Copy Fast, Graphic Speed, and Free Writing. These are briefly described in Table 1. There are separate sentences for copying for a younger (8–16 years) and older (17–25 years) age group. There is also an alternate Free Writing topic, appropriate for all ages. Within each age group, alternate versions of the sentence for copying and the Free Writing topic were administered to approximately half of the examinees. Due to the use of these different stimulus materials across participants it was not appropriate to use raw scores, so scaled scores were used in all analyses.

There were 112 examiners across all regions, who completed the assessment sessions between August 2021 and February 2023. Full details of testing are described in the DASH-2 test manual (Barnett et al., 2024). All testers were appropriately qualified, and extensive training on the administration and scoring of the test was provided to ensure consistency. Testing on DASH-2 was part of a larger project to standardize the Movement ABC-3 Test (Henderson & Barnett, 2023a) and Checklist (Henderson & Barnett, 2023b).

Data Preparation

Detailed information on the demographics for the UK and Australia/New Zealand samples taking part in the standardization of DASH-2 are presented in the test manual. Data analyses on the effect of specific demographic variables on test scores are also reported in the test manual. To justify the combination of data from the different countries for the development and analyses of test scores, two approaches were taken. First, to examine any effect of country (UK vs Australia/New Zealand) on test scores, analyses of variance were performed. A lack of significant differences along with small effect sizes on all comparisons provided evidence to support the combination of data across regions. Second, the measurement invariance between the UK and Australia/New Zealand was tested. The configural and metric invariances that were obtained supported the view that the model being tested was equivalent across countries.

Using the combined sample, developmental trends across age were examined for all five tasks (see Chapter 7 of the test manual). As in earlier editions, a steady increase in performance for all tasks was observed up to the age of 16 followed by a levelling off thereafter. These data supported the further use of age 16 as a dividing line between younger and older individuals for the production of different sentences to be copied and the subsequent measurement invariance analyses.

Of the 762 assessments conducted for the standardization project, 735 contained no missing scores. However, the 27 assessments for which a score was missing were still included in the analysis using a full information maximum likelihood estimator (Enders & Bandalos, 2001). For the Graphic Speed task, there was missing data for 2.2% of the sample. For all other tasks there was less than 1% missing data.

Results

Intercorrelations Between Task Scores

Scaled scores were used for the main data analyses. Prior to conducting factor analyses, it is important to ensure that the scores to be entered meet minimum requirements in relation to their correlations with each other. Pearson correlations were conducted between the individual task scaled scores for the DASH-2 standardization sample as a whole, for the younger and older age groups (8–16 years and 17–25 years) and for females and males separately.

Since all three correlation matrices were very similar, only the intercorrelations between task scores for the whole sample are presented in Table 3. All correlations were significant at $p < .001$. Correlations with Graphic Speed were low, ranging from .30 to .37. All other correlations between tasks were moderate, ranging from .52 (for the correlation between Free Writing and Alphabet Writing) to .77 (for the correlation between Copy Best and Copy Fast).

Factor Structure of the DASH-2 Scores

M Plus version 8.7 (Muthén & Muthén, 2017) was used for all confirmatory factor analysis models. As the data did not present any multivariate normality issue, the maximum likelihood

Table 3. Pearson intercorrelations between the task scores for the whole sample (N = 762)

	Copy best	Copy fast	Alphabet writing	Free writing
Copy Best	-	-	-	-
Copy Fast	.77**	-	-	-
Alphabet Writing	.56**	.60**	-	
Free Writing	.60**	.72**	.52**	-
Graphic Speed	.30**	.37**	.34**	.33**

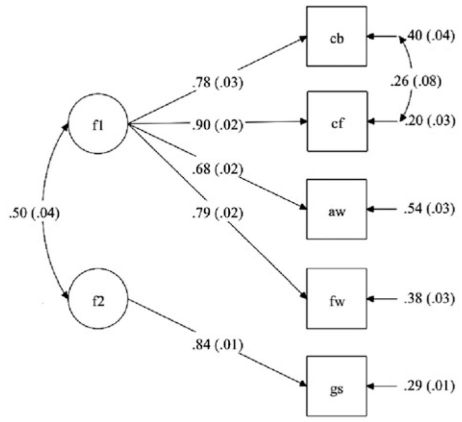
**p < .001.

estimator was used. The model tested is shown in Figure 1. This comprises a writing ability factor (f1) underlying the performance on the four core tasks: Copy Best, Alphabet Writing, Copy Fast, and Free Writing. The Graphic Speed task was removed from the main factor, but is included in the model as a grapho-motor factor (f2) in order to assess the relationship with the main factor. The model also allows for covariance between the error terms for the Copy Best and Copy Fast tasks, as they both use the same sentence for copying.

The standardized loadings are indicators of the relationship between the tasks and the factor underlying them and their squares are the percentages of variance taken into account by the factor. The residual variances are the remaining part of the score variance not taken into account by the factor $(1-\lambda^2)$.

The fit indices for this model, together with factor loadings and residual variances, are presented in Table 4. There is a significant correlation between the two factors of .50. Of the four tasks included in the main factor, Copy Fast has the highest saturation (.90), followed by Copy Best (.78) and Free Writing (.79). Alphabet Writing is the task with the lowest saturation (.68). The correlated errors between Copy Best and Copy Fast reflect the shared material of the sentences copied in each task. The introduction of these correlated residuals improved the model significantly, with a chi-square difference of 7.29 ($p < .01$).

A range of fit indices were computed. The model Chi-squared assesses the discrepancy between the observed and model fitted covariance matrices. Although it is recommended that the chi-squared value be greater than .05, in practice this criterion is often not met due to sample size and the demanding conditions of structural models. The Comparative Fit Index (CFI) and the Tucker Lewis Index (TLI), both less sensitive to sample size, compare the fit of the model to the fit of a baseline model (a model where task scores would not be underpinned by a common factor). CFI and TLI should be >.95 (Hu & Bentler, 1999) or .90 (Byrne, 1994). The Root Mean Square Error of Approximation (RMSEA) is a parsimony index, where lower values indicate better fit. Some authors suggest RMSEA should be <.08 (e.g. Awang, 2012) others <.05 (e.g. Byrne, 1994). In Table 4, the 90% RMSEA confidence interval is shown in brackets as well as the probability that the RMSEA is lower or equal to 0.05 and thus should be not significant (>.05). The Standardized



Note: cb – Copy Best, aw – Alphabet Writing, cf – Copy Fast, fw – Free Writing, gs – Graphic Speed.

Figure 1. Diagram of the model

Table 4. Fit indices of the model

Model	Chi-square	df	p-value	CFI	TLI	RMSEA (90% CI)	PCLOSE	SRMR
	18.37	4	<.01	.99	.98	.07 (.04–.10)	.138	.02

Root Mean Square Residual (SRMR) represents the square-root of the difference between the residuals of the observed and the model fitted covariance matrix. The SRMR should be <.08 (Byrne, 1994). The fit indices, shown in Table 4, were all considered to be satisfactory. Omega (ω), an estimation of the factor reliability (McDonald, 1999), was .86 for f1 and .71 for f2. The standardized loadings, standardized effect sizes, and R-squared are shown in Table 5. These measures indicate the strength and significance of the model.

Measurement Invariance

To assess the extent to which the factor structure is equivalent across two age ranges and for males and females, separately, Measurement Invariance analyses were performed using the Van de Schoot et al. (2012) procedure. This is a step-by-step procedure in which constraints on different parameters are progressively added to the model. The central idea is that by constraining these parameters to be equal across groups, the fit indices should not be significantly degraded. The three steps, Configural, Metric, and Scalar, are described below.

In the Configural invariant model, only the factor structure is identical. Factor loadings and intercepts can vary from group to group, regardless of how the groups are defined. This means that the same latent variables are measured in both groups, but the way in which they are measured may still differ. This level of invariance assesses the extent to which the same measurement model is suitable for describing the underlying structure of the two groups. If configural invariance is not found, the factor structure that underlies the scores is not the same for different groups and score configurations should therefore not be interpreted in the same way.

In the Metric invariant model, the factor structure is still identical but factor loadings are constrained to be equal between groups, in this case, age group or sex. This means that latent variables share the same scale across groups. The relationship between the latent variables and their indicators (the DASH-2 tasks) are, therefore, similar in both groups. If configural invariance is verified but metric invariance is not found, the relationship between one or several scores with their assigned factor is not the same in both groups and score configurations should therefore not be interpreted in the same way too.

In the Scalar invariant model, in addition to loadings, intercepts are also equated across groups. Intercepts are the score of indicators when the latent variable score is arbitrarily set to 0. This level of invariance makes it possible to compare the means of the latent variables between groups since the same observed score corresponds to the same level of ability. If scalar invariance is not found,

Table 5. Standardized loadings and effect sizes, and R-squared for the confirmatory factor analysis

Task	Standardized loadings	Standardized effect sizes	R-squared
Copy Best	0.78	0.03	0.60
Alphabet Writing	0.68	0.02	0.46
Copy Fast	0.90	0.02	0.81
Free Writing	0.79	0.02	0.63
Graphic Speed	0.84	0.01	0.71

any differences found between groups are not necessarily related to the latent trait, but to the measurement non-equivalence – the instrument favours a group despite an equivalent level of ability. Conversely, if scalar invariance is found, differences in latent variable means can be interpreted as reflecting true differences between groups.

The level of measurement invariance is typically evaluated by changes in the CFI, RMSEA, and SRMR. Measurement invariance is supported if the change in the CFI (ΔCFI) of two models is 0.01 or less (Chen, 2007; Cheung & Rensvold, 2002) and the changes in the SRMR (ΔSRMR) less than 0.03 or RMSEA (ΔRMSEA) less than 0.015 (Chen, 2007).

Table 6 presents the fit indices of the different Measurement Invariance models across two age group (8–16 years and 17–25 years) models and Table 7 the differences between these models. As demonstrated in Tables 6 and 7, the configural model shows an adequate fit, indicating that the latent structure is consistent across the two age groups. This finding then allows for the application of the metric and scalar models to evaluate incremental fit differences. Both the metric and scalar models also display satisfactory fit indices. Transitioning from the configural model to the metric invariance model results in negligible changes in CFI, SRMR, and RMSEA (ΔCFI = .00, ΔSRMR = .00, ΔRMSEA = .00), all within acceptable limits, confirming the equivalence of factor loadings between the age groups. Subsequently, comparing the scalar invariance model with the metric invariance model shows negligible changes in CFI and SRMR (ΔCFI = 0.00, ΔSRMR = .00) and a minimal change in RMSEA (ΔRMSEA = −.01), indicating intercept equivalence.

Table 8 presents the fit indices of the different Measurement Invariance models across males and females and Table 9 the differences between these models. As these tables show, the configural model exhibits adequate fit, indicating that the latent structure is equivalent in males and females. Subsequently, the metric and scalar models can therefore be evaluated to assess incremental fit differences. The metric and scalar models also show adequate fit indices. Switching from the configural model to the metric invariance model results in negligible variations in CFI and minimal variations in SRMR and RMSEA (ΔCFI = .00, ΔSRMR = −.01, ΔRMSEA = .01). These variations are all within acceptable limits, confirming the equivalence of factor loadings between the sexes. Next, the scalar invariance model was compared with the metric invariance model, with negligible changes in fit indices (ΔCFI = .00, ΔSRMR = .00, ΔRMSEA = .00), indicating intercept equivalence.

Discussion

As an important component of literacy, handwriting continues be part of the school curriculum in many countries and contributes to learning and attainment throughout education. Without support and appropriate accommodations, students with handwriting difficulties may struggle to keep up with writing demands and to demonstrate their knowledge and understanding. As a result, they may under-perform in class activities and written examinations (Graham, 2022).

Since first being developed, the DASH and DASH17+ have been widely used in clinics, schools, and further/higher education settings to help identify those with handwriting difficulties, plan appropriate support, and to use as evidence in applying for accommodations and/or additional

Table 6. Measurement Invariance model fit indices for age group

Model	chisq	df	p	CFI	TLI	RMSEA	PCLOSE	SRMR
Configural	19.093	8	.014	.99	.98	.06 (.03–.10)	.272	.02
Metric	26.534	11	.005	.99	.98	.06 (.03–.09)	.242	.04
Scalar	27.673	14	.016	.99	.99	.05 (.02–.08)	.446	.04

Table 7. Comparison of Measurement Invariance models for age group

Model	Δ chisq	Δ df	p	Δ CFI	Δ TLI	Δ RMSEA	Δ SRMR
Configural vs Metric	7.441	3	.059	.00	.00	.00	.02
Configural vs Scalar	8.58	6	.199	.00	.01	−.01	.02
Metric vs Scalar	1.139	3	.768	.00	.01	−.01	.00

resources. Their contribution to the assessment process has also been recognized through various professional bodies including the Specific Learning Difficulties (SpLD) Assessment Standards Committee (SASC, 2024) in the UK and the Western Australian Occupational Therapy Association (DOT(WA), 2019). Both editions of the test have also been used in research studies to examine handwriting in various neurodevelopmental disorders including dysgraphia (Breaux & Munsell, 2025a; Richards et al., 2015), developmental coordination disorder (DCD) (Prunty & Barnett, 2020), and dyslexia (Sumner et al., 2013).

The DASH-2 can be used in different ways, depending on the purpose of assessment. In some situations, it is appropriate to use the DASH-2 in isolation. For example it can be used in a group setting to screen for handwriting difficulties on school entry; in this case the main focus may be just on the Total Standard Score. In contrast, when used in an individual setting, a range of information can be gathered to help understand the nature of a handwriting difficulty and help plan an intervention programme. All three editions of DASH differ from other handwriting tests in that they offer a multi-faceted measure of the speed of handwriting derived from four separate but related tasks. These differ in length and in the task demands, reflecting the complex processes involved in handwriting and writing. When inspected separately, as a profile, or a total score, they not only help a tester make decisions about diagnosis or service provision, they also help to pinpoint specific challenges the student is encountering. For example, Breaux and Munsell (2025a) draw particular attention to the value of comparing the Copy Best and Copy Fast tasks when differentiating different types of dysgraphia. Although described in the test manual as a supplementary score, the standard score for Graphic Speed, reflecting the ability to control a writing implement, is an important feature of the test. Performance on this task can lead assessors to better understand the possible contribution of motor difficulties to handwriting performance and consider the possibility of a broader motor difficulty in some students. In addition to the quantitative scores, a range of qualitative information can be gathered from observation of the student writing and later, through close inspection of the script. It can also be useful to administer other assessments alongside DASH-2 as part of a broader writing assessment. This may include measures of handwriting legibility, spelling, and written expression (Barnett & Prunty, 2021).

The latest edition, DASH-2, provides extended norms from 8 to 25 years. Data from the UK, Australia, and New Zealand are comparable, indicating that the norms can be used with confidence across these three regions. Furthermore, Breaux and Munsell (2025b) report a study comparing mean performance on the DASH-2 between a sample in the United States of America (USA) and a matched control group from the DASH-2 normative sample. There was no significant group

Table 8. Measurement Invariance model fit indices for sex

Model	chisq	df	p	CFI	TLI	RMSEA	PCLOSE	SRMR
Configural	22.772	8	.004	.99	.98	.07 (.04–.10)	.146	.02
Metric	24.167	11	.012	.99	.99	.06 (.03–.09)	.332	.03
Scalar	30.948	14	.006	.99	.99	.06 (.03–.08)	.315	.03

Table 9. Comparison of Measurement Invariance models for sex

Model	Δ chisq	Δ df	p	Δ CFI	Δ TLI	Δ RMSEA	Δ SRMR
Configural vs metric	1.395	3	.707	.00	.01	−.01	.01
Configural vs scalar	8.176	6	.225	.00	.01	−.01	.01
Metric vs scalar	6.781	3	.079	.00	.00	.00	.00

difference for the individual tasks or for the total scores, suggesting that the test is also suitable for use in the USA.

Analyses of the normative data for DASH and subsequent editions have supported the calculation of a total standard score from the four core tasks: Copy Best, Alphabet Writing, Copy Fast, and Free Writing as a reliable global measure of handwriting speed across different writing tasks and under different conditions. The Graphic Speed task then acts as a valuable supplementary measure of grapho-motor ability, independent of a language element. In the current test manual, Confirmatory Factor Analysis (CFA) supporting this structure is reported. Two models were initially tested using CFA. The first included all five tasks in a common factor model and the second excluded Graphic Speed. Although the indices showed a good fit for both models, as previously reported for DASH and DASH17+, Graphic Speed had a smaller loading on the common factor than the other tasks (0.40 versus 0.66 to 0.93) and further supported presenting the DASH-2 scores as comprising two components, one focussing on tasks which come under the umbrella of the broader concept of writing, the other providing a measure of the physical ability to control a writing instrument.

The present study included further investigation of the structure of the test using CFA, this time taking account of the fact that the stimulus material for two of the tasks, Copy Best and Copy Fast is identical. The introduction of these correlated residuals improved the model significantly. The analyses were also extended to investigate measurement invariance across age group and sex nested within the broader target population. This is an important aspect of construct validation, as it considers whether the test's factor structure, factor loadings, and intercepts are equivalent across different groups when subjected to increasingly restrictive parameter constraints.

As before, the CFA replicated the previous finding revealing one main factor comprising the four core tasks, with factor loadings on the main writing factor, ranging from the highest at .90 for Copy Fast to the lowest at .68 for Alphabet Writing and a second grapho-motor factor, comprising only the Graphic Speed task which has a substantial loading of 0.84. In addition, however, a significant correlation between the main writing factor and the grapho-motor speed factor suggested they are not entirely independent of each other. While we consider that the fit indices for the model tested in this study strongly support the current presentation of the test, this does not preclude the possibility that alternate models could be tested. The model of DASH-2 presented in the CFA was used to investigate measurement invariance (MI) across age groups and sex to determine whether the test scores can be confidently interpreted across groups in the same way. This is the first study to scrutinize the DASH-2 (or previous versions of the test) in this way and is important because without evidence of measurement invariance (MI) across different demographic groups, the extent to which the instrument may be susceptible to measurement error is not known. Three levels of MI were investigated: configural, metric, and scalar. Configural invariance determines whether the pattern of the factor loadings is the same across groups. Metric invariance considers both the factor structure and the factor loadings, indicating whether the importance of each task in relation to the factors is the same across groups. Finally, scalar invariance considers whether the way the test measures the level of ability is the same across groups. The validation of these three invariance models confirmed that the factor model of the DASH-2 maintained strong measurement invariance across the two age groups and across males and females. These findings indicate that test users can be confident that the scores obtained on DASH-2 are not influenced by measurement error due to age or sex and that the structure of the test is reliable. With regard to interpretation, therefore, the findings also mean that users can be confident that differences in scaled scores actually do reflect differences in the construct being measured rather than any differences in the way the test functions across these groups.

Conclusion and Implications for Practice

The findings of this study add substantially to the evidence of construct validity presented for DASH-2 to date and have important implications for test users. DASH-2 is unique in that it offers a multi-faceted measure of speed of handwriting. The new data reported in this study, demonstrating invariance across age group and sex, means that clinicians and educators can interpret group differences as reflecting actual differences in handwriting speed rather than measurement error. Most importantly, this supports the fairness of using DASH-2 scores to make diagnostic or accommodation decisions across different age levels and across males and females.

In addition to the main scores examined in this study, DASH-2 provides a range of quantitative and qualitative information useful for the practitioner. This includes a profile of performance across the five tasks plus a measure of the difference between the two copying tasks (reflecting the ability to speed up on instruction). DASH-2 also provides an opportunity to record the process of writing through qualitative observations made during task performance (e.g. posture and how letters are formed). Close inspection of the final product in all tasks can also provide detailed information useful for intervention planning. When taken together with the standard test scores, DASH-2 therefore allows testers from many different professions to produce a comprehensive picture of an individual's handwriting competence.

ORCID iD

Anna L. Barnett https://orcid.org/0000-0003-3800-9468

Ethical Consideration

Ethical approval was granted from Pearson Clinical Scientific Council, Texas USA.

Consent to Participate

Informed consent was provided by participants and/or their parent/caregiver.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The project was funded by Pearson Clinical.

Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: ALB, BS, and SEH receive royalties from the sales of DASH-2. LCV is employed by Pearson Clinical.

Data Availability Statement

The data are not made available.

References

Ahmed, Y., Kent, S., Cirino, P. T., & Keller-Margulis, M. (2022). The not-so-simple view of writing in struggling readers/writers. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, *38*(3), 272–296. https://doi.org/10.1080/10573569.2021.1948374

Alves, R. A., Limpo, T., Fidalgo, R., Carvalhais, L., Pereira, L. Á., & Castro, S. L. (2016). The impact of promoting transcription on early text production: Effects on bursts and pauses, levels of written language, and writing performance. *Journal of Educational Psychology*, *108*(5), 665–679. https://doi.org/10.1037/edu0000089

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed. text rev.). APA.

Araújo, S., Domingues, M., & Fernandes, T. (2022). From hand to eye: A meta-analysis of the benefit from handwriting training in visual graph recognition. *Educational Psychology Review*, *34*(3), 1577–1612. https://doi.org/10.1007/s10648-021-09651-4

Australian Curriculum, Assessment and Reporting Authority (ACARA) (2022). *Australian Curriculum. English F-10 Version 9.0. Curriculum content F-6*.

Awang, Z. (2012). A handbook on SEM. Structural Equation Modeling.

Barnett, A., Henderson, S. E., & Scheib, B. (2024). *The Detailed Assessment of the Speed of Handwriting - 2nd edition (DASH-2)*. Pearson.

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

Barnett, A., Henderson, S. E., Scheib, B., & Schulz, J. (2010). *The Detailed Assessment of Speed of Handwriting 17+ (DASH 17+)*. Pearson Assessment.

Barnett, A. L., Henderson, S. E., Scheib, B., & Schulz, J. (2009). Development and standardisation of a new handwriting speed test: The DASH. *British Journal of Educational Psychology Monograph Series II*, *1*(1), 137–157. https://doi.org/10.1348/000709909X421937

Barnett, A. L., Henderson, S. E., Scheib, B., & Schulz, J. (2011). Handwriting difficulties and their assessment in young adults with DCD: Extension of the DASH for 17-25 year olds. *Journal of Adult Development*, *18*(3), 114–121. https://doi.org/10.1007/s10804-011-9121-3

Barnett, A. L., & Prunty, M. (2021). Handwriting difficulties in Developmental Coordination Disorder (DCD). *Current Developmental Disorders Reports*, *8*(1), 6–14. https://doi.org/10.1007/s40474-020-00216-8

Bazerman, c., Graham, S., Applebee, A. N., Matsuda, P. K., Berninger, V. W., Murphy, S., Brandt, D., Wells Rowe, D., & Schleppegrell, M. (2017). Taking the long view on writing development. *Research in the Teaching of English*, *51*(3), 351–360. https://doi.org/10.58680/rte201728980

Berninger, V. W., & Winn, W. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). Guilford.

Breaux, K. C., & Munsell, E. (2025a). *Using the DASH-2 to identify dysgraphia*. NCS Pearson. [Technical Report].

Breaux, K. C., & Munsell, E. (2025b). *Regional comparison study using the DASH-2*. NCS Pearson. [Technical Report].

Brossard-Racine, M., Shevell, M., Snider, L., Belanger, S. A., Julien, M., & Majnemer, A. (2015). Persistent handwriting difficulties in children with ADHD after treatment with stimulant medication. *Journal of Attention Disorders*, *19*(7), 620–629. https://doi.org/10.1177/1087054712461936

Byrne, B. M. (1994). *Structural equation modelling with EQS and EQS/Windows*. Sage Publications.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/s15328007sem0902_5

Chung, P. J., Patel, D. R., & Nizami, I. (2020). Disorder of written expression and dysgraphia: Definition, diagnosis, and management. *Translational Pediatrics*, *9*(Suppl 1), S46–S54. https://doi.org/10.21037/tp.2019.11.01

Coombe, G., Lester, A., & Moores, L. (2020). Online and on-screen assessment in high-stakes, sessional qualifications: A review of the barriers to greater adoption and how these might be overcome. *Ofqual*. Ofqual/20/6723/1.

Department for Education (DfE). (2013). The national curriculum in England. Key stages 1 and 2 framework document. https://assets.publishing.service.gov.uk/media/5a81a9abe5274a2e8ab55319/PRIMARY_national_curriculum.pdf

Developmental Occupational Therapy (WA) Inc. (DOT(WA)). (2019). Motor dysgraphia: Clinical practice guidelines for occupational therapists in Western Australia. https://dotwa.org.au/v2/wp-content/uploads/2020/06/DOT10758-MOTOR-DYSGRAPHIA-Clinical-Guidelines_v51.pdf

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 430–457. https://doi.org/10.1207/s15328007sem0803_5

Engel-Yeger, B., Nagauker-Yanuv, L., & Rosenblum, S. (2009). Handwriting performance, self-reports, and perceived self-efficacy among children with dysgraphia. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *63*(2), 182–192. https://doi.org/10.5014/ajot.63.2.182

Feng, L., Lindner, A., Ji, X. R., & Malatesha Joshi, R. (2019). The roles of handwriting and keyboarding in writing: A meta-analytic review. *Reading and Writing*, *32*(1), 33–63. https://doi.org/10.1007/s11145-017-9749-x

Gargot, T., Asselborn, T., Pellerin, H., Zammouri, I., Anzalone, S., Casteran, L., Dillenbourg, P., Cohen, D., & Jolly, C. (2020). Acquisition of handwriting in children with and without dysgraphia: A computational approach. *PloS One*, *15*(9), Article e0237575. https://doi.org/10.1371/journal.pone.0237575

Gosse, C., Parmentier, M., & Van Reybroeck, M. (2021). How do spelling, handwriting speed, and handwriting quality develop during primary school? cross-classified growth curve analysis of children's writing development. *Frontiers in Psychology*, *12*, Article 685681. https://doi.org/10.3389/fpsyg.2021.685681

Graham, S. (2022). A walk through the landscape of writing: Insights from a program of writing research. *Educational Psychologist*, *57*(2), 55–72. https://doi.org/10.1080/00461520.2021.1951734

Graham, S., Fishman, E. J., Reid, R., & Hebert, M. (2016). Writing characteristics of students with attention deficit hyperactive disorder: A meta-analysis. *Learning Disabilities Research & Practice*, *31*(2), 75–89. https://doi.org/10.1111/ldrp.12099

Graham, S., Harris, K. R., & Hebert, M. (2011). Presentation effects in scoring writing. *Focus on Exceptional Children*, *44*(4), 1–12. https://doi.org/10.17161/fec.v44i4.6687

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, *81*(4), 710–744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Kavanaugh, C., & Talukdar, J. (2018). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research*, *88*(2), 243–284. https://doi.org/10.3102/0034654317746927

Hebert, M., Kearns, D. M., Hayes, J. B., Bazis, P., & Cooper, S. (2018). Why children with dyslexia struggle with writing and how to help them. *Language, Speech, and Hearing Services in Schools*, *49*(4), 843–863. https://doi.org/10.1044/2018_LSHSS-DYSLC-18-0024

Henderson, S. E., & Barnett, A. L. (2023a). *Movement Assessment Battery for children* - (3rd edition (Movement ABC-3). Test manual. Pearson.

Henderson, S. E., & Barnett, A. L. (2023b). *Movement Assessment Battery for Children* - (3rd edition (Movement ABC-3). Checklist manual. Pearson.

Hen-Herbst, L., & Rosenblum, S. (2022). Handwriting and motor-related daily performance among adolescents with dysgraphia and their impact on physical health-related quality of life. *Children*, *9*(10), 1437. https://doi.org/10.3390/children9101437

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kandel, S., Peereman, R., Grosjacques, G., & Fayol, M. (2011). For a psycholinguistic model of handwriting production: Testing the syllable-bigram controversy. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(4), 1310–1322. https://doi.org/10.1037/a0023094

Killeen, H., Dempsey, M., & O'Mahony, P. (2006). The Irish adaptation of the handwriting speed test. *Dublin, Ireland: The Association of Occupational Therapists of Ireland*.

McCloskey, M., & Rapp, B. (2018). Developmental dysgraphia: An overview and framework for research. In *Developmental dysgraphia*. Routledge.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum.

Medwell, J., Strand, S., & Wray, D. (2009). The links between handwriting and composing for Y6 children. *Cambridge Journal of Education*, *39*(3), 329–344. https://doi.org/10.1080/03057640903103728

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.) Muthén & Muthén.

Pritchard, V. E., Malone, S. A., & Hulme, C. (2021). Early handwriting ability predicts the growth of children's spelling, but not reading, skills. *Scientific Studies of Reading*, *25*(4), 304–318. https://doi.org/10.1080/10888438.2020.1778705

Prunty, M. M., & Barnett, A. L. (2020). Accuracy and consistency of letter formation in children with Developmental Coordination Disorder. *Journal of Learning Disabilities*, *53*(2), 120–130. https://doi.org/10.1177/0022219419892851

Prunty, M. M., Barnett, A. L., Wilmut, K., & Plumb, M. S. (2014). An examination of writing pauses in the handwriting of children with Developmental Coordination Disorder. *Research in Developmental Disabilities*, *35*(11), 2894–2905. https://doi.org/10.1016/j.ridd.2014.07.033

Richards, T. L., Grabowski, T. J., Boord, P., Yagle, K., Askren, M., Mestre, Z., Robinson, P., Welker, O., Gulliford, D., Nagy, W., & Berninger, V. (2015). Contrasting brain patterns of writing-related DTI parameters, fMRI connectivity, and DTI–fMRI connectivity correlations in children with and without dysgraphia or dyslexia. *NeuroImage: Clinical*, *8*, 408–421. https://doi.org/10.1016/j.nicl.2015.03.018

Salas, N., & Silvente, S. (2020). The role of executive functions and transcription skills in writing: A cross-sectional study across 7 years of schooling. *Reading and Writing*, *33*(4), 877–905. https://doi.org/10.1007/s11145-019-09979-y

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

Specific Learning Difficulties (SpLD). (2024). Assessment Standards Committee (SASC). https://www.sasc.org.uk/media/j2uf115h/dash2-guidance-stec-september-2024.pdf

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

Sumner, E., Connelly, V., & Barnett, A. L. (2014). The influence of spelling ability on handwriting production: Children with and without dyslexia. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(5), 1441–1447. https://doi.org/10.1037/a0035785

Thibon, L. S., Gerber, S., & Kandel, S. (2018). The elaboration of motor programs for the automation of letter production. *Acta Psychologica*, *182*, 200–211. https://doi.org/10.1016/j.actpsy.2017.12.001

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

Van Galen, G. P. (1991). Handwriting: Issues for a psychomotor theory. *Human Movement Science*, *10*(2-3), 165–191. https://doi.org/10.1016/0167-9457(91)90003-g

Wiley, R. W., & Rapp, B. (2021). The effects of handwriting experience on literacy learning. *Psychological Science*, *32*(7), 1086–1103. https://doi.org/10.1177/0956797621993111

Zwicker, J. G., Suto, M., Harris, S. R., Vlasakova, N., & Missiuna, C. (2017). Developmental coordination disorder is more than a motor problem: Children describe the impact of daily struggles on their quality of life. *British Journal of Occupational Therapy*, *81*(2), 65–73. https://doi.org/10.1177/0308022617735046