



## The Open Journal of Occupational Therapy

Volume 7  
Issue 4 Fall 2019

Article 2


October 2019

### Handwriting Performance of Typical Second-Grade Students as Measured by the Evaluation Tool of Children's Handwriting - Manuscript and Teacher Perceptions of Legibility

Diane M. Long  
*Ithaca College*, [dlong@ithaca.edu](mailto:dlong@ithaca.edu)

James Conklin  
*Ithaca College*, [conklin@ithaca.edu](mailto:conklin@ithaca.edu)

Follow this and additional works at: <https://scholarworks.wmich.edu/ojot>

 Part of the Educational Assessment, Evaluation, and Research Commons, and the Occupational Therapy Commons

#### Recommended Citation

Long, D. M., & Conklin, J. (2019). Handwriting Performance of Typical Second-Grade Students as Measured by the Evaluation Tool of Children's Handwriting - Manuscript and Teacher Perceptions of Legibility. *The Open Journal of Occupational Therapy*, 7(4), 1-14. <https://doi.org/10.15453/2168-6408.1492>

This document has been accepted for inclusion in The Open Journal of Occupational Therapy by the editors. Free, open access is provided by ScholarWorks at WMU. For more information, please contact [wmu-scholarworks@wmich.edu](mailto:wmu-scholarworks@wmich.edu).

---

# Handwriting Performance of Typical Second-Grade Students as Measured by the Evaluation Tool of Children's Handwriting - Manuscript and Teacher Perceptions of Legibility

## Abstract

*Background:* The purpose of the study was to describe scores achieved by typical second-grade students on the Evaluation Tool of Children's Handwriting – Manuscript and to compare scores with teacher perceptions.

*Method:* As part of a larger study, the ETCH-M was administered to 74 second-grade students. Teachers scored classroom samples of handwriting assignments using a researcher-developed scale and scores were compared to ETCH-M scores to determine cutoff values for good versus poor handwriting.

*Results:* Mean scores for total word legibility, total letter legibility, and total numeral legibility were 88.82%, 84.30%, and 89.26%, respectively. Cutoff scores below 82% for word legibility and 77% for letter legibility for second-grade students based on teacher perceptions of below average handwriting are cautiously suggested. Research with a larger dataset is needed. Boys scored significantly lower on the ETCH-M and this finding warrants further research.

*Conclusion:* The findings add to the limited body of information about the psychometric properties of the ETCH-M and the normative performance of typical second-grade students.

## Comments

The authors report that they have no conflicts of interest to disclose.

## Keywords

handwriting, assessment, normative data, school-based practice

## Cover Page Footnote

The authors would like to acknowledge graduate students who assisted in data collection and scoring aspects of this study.

## Credentials Display

Diane M. Long, Ed.D., MOTR/L

James Conklin, Ph.D.

Copyright transfer agreements are not obtained by The Open Journal of Occupational Therapy (OJOT). Reprint permission for this Applied Research should be obtained from the corresponding author(s). Click here to view our open access statement regarding user rights and distribution of this Applied Research.

DOI: 10.15453/2168-6408.1492

Occupational therapy is intended to enable “participation in roles, habits, and routines in home, school, workplace, community, and other settings” (American Occupational Therapy Association [AOTA], 2014, p. S1). A primary occupation for students in school is communicating through handwriting, and difficulties with production of legible handwriting often result in a referral to occupational therapy (Feder, Majnemer, & Synnes, 2000; Hammerschmidt & Sudsawad, 2004; Woodward & Swinth, 2002).

Occupational therapy services are provided as a related service in a variety of ways to meet the intent of the 2004 Individuals with Disabilities Education Improvement Act. Direct one-on-one intervention outside or inside of the classroom (Donica, Larson, & Zinn, 2012), consultation (Donica, 2015), integrative co-teaching methods (Case-Smith, Holland, & Bishop, 2011; Case-Smith, Holland, Lane, & White, 2012; Gerde, Foster, & Skibbe, 2014), and the Response to Intervention (RtI) plans (Cahill, McGuire, Krundick, & Lee, 2014) have all been described in the literature.

Reported outcomes of occupational therapy intervention for handwriting suggest that consultative (Donica, 2015), integrative (Case-Smith et al., 2011), and co-teaching (Case-Smith et al., 2012) approaches using a variety of handwriting interventions are effective in improving the handwriting skills of students in schools. Hoy, Egan, and Feder (2011) reported that interventions must include more than 20 practice sessions to impact handwriting improvement. There is also support in the literature for occupational therapy handwriting intervention (Clark, Brouwer, Schmidt, & Alexander, 2008; Reeder, Arnold, Jeffries, & McEwen, 2011; Zylstra & Pfeiffer, 2016) or precursor skills (Ohl et al., 2013), specifically through the RtI process in kindergarten through the third grade (Cahill et al., 2014).

Regardless of the form of the intervention, there must be some evidence to support both the need for and the outcomes of school-based intervention when services are provided (<https://www.copyright.gov/legislation/pl108-446.pdf>). There are many handwriting assessments available (Feder & Majnemer, 2003) with the potential to meet the needs for a data-driven justification for intervention in handwriting. Therapists must decide the best choice of assessment, as the assessments have different properties and potentially serve different purposes when reporting on student skill levels. Tools that are norm-referenced, such as the Test of Handwriting Skills (Milone, 2007) and the Minnesota Handwriting Assessment (Reisman, 1999), allow comparison against a specified population and are useful for the purposes of diagnosis (Portney & Watkins, 2009). Criterion-referenced tools yield information about an individual's performance along a continuum in relation to an absolute standard or benchmark but do not compare the individual's performance to any specific population (Portney & Watkins, 2009). For instance, scores on The Print Tool (Olsen & Knapton, 2016) can be compared to suggested grade-level benchmarks, and scores on the Evaluation Tool of Children's Handwriting - Manuscript (ETCH-M) (Amundson, 1995) are reported out of 100% legibility. In addition, therapist-designed tools are frequently used that have not been subjected to psychometric analysis. In school-based practice, normative comparison with other students in the grade level or comparison with an absolute standard might be necessary to justify the need for occupational therapy services.

### **Psychometric Properties of the ETCH-M**

The ETCH-M was developed as a measure of global legibility and writing speed and includes several components of legibility appropriate for use with children who have mild motor and learning delays (Amundson, 1995). The ETCH-M consists of seven subtests. Three subtests require writing from memory: upper case letters, lower case letters, and numerals one through 12. Two subtests require

copying five-word sentences: one from a page on the student's desk (near point copying) and one from a sample at a prescribed distance (far point copying). There is an oral dictation subtest of two nonsense words and a zip code and a spontaneous sentence writing subtest. Percentage scores of the number of legible characters or words are calculated per subtest and then totaled per categories of letters, words, or numerals (Amundson, 1995).

Some psychometric properties of the ETCH-M have been described. When initially published, Amundson (1995) included an analysis of inter-rater reliability in the administration manual and suggested that the total scores for word, letter, and numeral legibility were less subject to variation than were individual subtests and that letter scores were more stable than word scores. Amundson reported differences in inter-rater reliability between experienced and inexperienced raters with intraclass correlation coefficients (ICCs) of .84 for total letters, .82 for total numbers, and .48 for total words. As per Amundson, content validity was supported with three prepublication pilot studies that were used to refine the ETCH-M, but data was not reported.

Diekema, Deitz, and Amundson (1998) conducted a study of the test-retest reliability of the ETCH-M and gleaned reliability ICCs of .77, .71, and .63 for total letter, total word, and total numeral legibility, respectively. The ICCs of the subtests, other than alphabet uppercase, were not deemed adequate for making clinical decisions about intervention (Diekema, Deitz, & Amundson, 1998). The total scores were within acceptable levels as compared to other handwriting assessments and the subjective nature of scoring, even though the test-retest reliability was lower than statistically ideal (Diekema et al., 1998). No other psychometric studies or normative data analyses were included in the publication manual or reported in the literature. Amundson (1995) suggested that further testing of the tool was warranted.

The grading of handwriting tends to be more subjective than other subjects in school. In elementary education, the grading of many subjects is based on relatively standard continuums (Portney & Watkins, 2009). Students are given numerical grades out of 100%, or student skills are expressed along a continuum, such as *exceeds expectations*, *meets expectations*, or *needs improvement*. Often, a grade of 70 is used as the absolute standard for passing or failing for knowledge in a subject area (Portney & Watkins, 2009) when objective information is tested (e.g., how to spell words, facts about history).

Handwriting legibility is less objective, especially when assessed from a global legibility perspective, and is the result of several aspects of executive functioning (Altemeier, Abbott, & Berninger, 2008; Rosenblum, 2018). The inherent complexity of handwriting proficiency makes it more important to consider legibility developmentally (Abbott & Berninger, 1993; Graham, Berninger, Weintraub, & Schafer, 1998) with a need for attention to grade level and typical performance. Given that there were not any large-scale normative studies for handwriting legibility as scored by the ETCH-M, the user must interpret the results in the absence of normative data or an absolute standard per grade level.

Some researchers have reported findings that begin to identify absolute standards for performance on the ETCH-M. Feder, Majnemer, Bourbonnais, Blayney, and Morin (2007) conducted a study of 69 first-grade students and reported mean scores of 67.8 ( $SD = 23.3$ ) for total word scores, 77.4 ( $SD = 13.7$ ) for total letter scores, and 86.9 ( $SD = 16.0$ ) for total numeral scores, with boys performing significantly lower than girls.

Brossard-Racine, Mazer, Julien, and Majnemer (2012) reported cutoff scores on the ETCH-M for 26 second-grade and third-grade students diagnosed with ADHD based on occupational therapists' perceptions (yes or no) of the need for occupational therapy services after viewing the ETCH-M samples. Total word legibility scores of 75% and total letter legibility scores of 76% were deemed appropriate cutoff scores for referral for occupational therapy services and changes of more than 10% and 6%, respectively, were suggested as minimally clinically important differences (Brossard-Racine et al., 2012).

The relationship between scores on the ETCH-M and teachers' perceptions of handwriting has been explored with mixed results (Feder, Majnemer, Bourbonnais, Blayney, & Morin, 2007; Grace-Fredrick, 1998; Sudsawad, Trombly, Henderson, & Tickle-Degnen, 2001). Feder et al. (2007) gathered teachers' perceptions of typical first-grade students' handwriting through use of a researcher-designed checklist. The teachers rated handwriting legibility and speed as "*above average, average, needs improvement or very poor*" (p. 48) and compared each student's handwriting to classroom peers. The teachers' scores correlated with the ETCH-M were reported as " $r = 0.40-0.45; p < 0.05$ " (p. 52).

In a study by Grace-Fredrick (1998) of 133 second-grade students, teachers responded to a 5-point rating scale for "overall printing ability" (p. 38) that ranged from *much less than average* (score of 1) to *much above average* (score of 5). Teacher judgments were based on "general knowledge of the child's handwriting and relative ranking within the class" (p. 17). Teachers' perceptions of average or below average correlated with total scores and eight of 11 subtests (Grace-Fredrick, 1998). Grace-Fredrick reported percentage mean scores of 90.32 ( $SD = 6.91$ ) for total letters, 82.62 ( $SD = 19.72$ ) for total words, and 95.04 ( $SD = 8.02$ ) for total numerals for students with less than average handwriting legibility as perceived by teachers. Both studies suggested modest correlations between the ETCH-M scores and teachers' perceptions.

In contrast with the studies by Feder et al. (2007) and Grace-Fredrick (1998) that included students with typical handwriting, Sudsawad, Trombly, Henderson, and Tickle-Degnen (2001) studied 45 first-grade students who had handwriting difficulties. Sudsawad et al. used six research-designed multiple choice questions to determine teachers' perceptions about students' handwriting in comparison to their peers. Questions asked about overall legibility and student skills similar to the subtests of the ETCH-M, such as copying from models and dictation. Teachers were asked to respond to each question based on a 7-point choice from *much below average* to *much above average*. The findings suggested no relationship between teachers' perceptions and scores on the ETCH-M for this group of students.

### **Purpose**

The psychometric properties of the ETCH-M have not been fully researched for second-grade students. A search of the literature yielded no other studies specific to second-grade students' performances on the ETCH-M. The purpose of this study was threefold. First, we wanted to describe the typical performance of second-grade students on the ETCH-M to add to the psychometric data reported in the literature. Normative data for second-grade students has not been reported for typically developing children and the information would be helpful to therapists who are making decisions about recommendations for therapeutic intervention. Second, we wanted to know if gender influenced scores on the ETCH-M, as prior research suggested gender effects on writing (Feder et al., 2007; Graham et al., 1998). Third, we wanted to compare teachers' perceptions of handwriting for second-grade students with the ETCH-M scores. Since the primary purpose of school-based intervention is to support mastery of occupations in the classroom environment, the relationship between teachers' perceptions and the

ETCH-M scores is important to know for purposes of ecological validity; previous studies reported mixed findings.

### **Method**

Approvals by appropriate school district personnel and the human subjects review board at the college were obtained. Schools and classroom teachers that participated in the study were recruited via personal contact using convenience sampling. All teachers who responded to invitations to participate were included, and all students in each teachers' classroom were invited to participate. The data reported herein were part of a larger study.

### **Participants**

Eight classrooms in four school districts in central New York participated. Five of the classrooms were in one small city school district classified as a high need district, and three classrooms were from faith-based schools in other districts. The number of participants per classroom ranged from four to 14.

The student participants were included in the study based on parental permission through informed consent forms sent home and collected by participating classroom teachers. The parents completed demographic information and consent forms for each student participant.

### **Procedures**

The ETCH-M was administered to second-grade students in the school environment over a 2-week period in early November. The criteria for inclusion were that students were educated in the regular second-grade classroom, understood spoken English, and could handwrite independently. The students who were receiving educational or therapeutic support services were included, as the ETCH-M is designed to be administered to children with known difficulties. The participants were administered the ETCH-M individually in locations outside the classrooms to reduce distractions. The students were familiar with the cafeterias and libraries used for testing. Testing was done during times that did not compete with typical uses of the spaces.

Graduate students were taught administration and scoring of the ETCH-M didactically by the primary researcher, completed the scoring practice tests included in the ETCH manual, and practiced scoring sets of completed samples until attainment of ICCs of at least .88 occurred for total word, total letter, and total numeral scores. All handwriting samples collected from the participants were first scored by graduate students. To maintain inter-rater reliability, every fifth handwriting sample was scored separately by graduate students who then compared their results and resolved any discrepancies. All ETCH-M samples were then master scored by the primary researcher, who was experienced in the administration and scoring of the ETCH-M. The master scoring was done to ensure scoring consistency across samples. In the small number of instances when scoring was corrected, the differences were predominately related to the subtest scores with minimal effects on the total scores; the master scoring was used for analysis.

The teachers were asked to grade the participants' handwriting based on a typical classroom assignment using a sliding scale of 0.0 to 5.0. The teachers were given a key that identified well below average for grade level as 0.0 - 1.0, below average as 1.1- 2.0, average as 2.1 - 3.0, above average as 3.1 - 4.0, and superior as 4.1 - 5.0. The teachers could choose any increment in the ranges. The teachers graded all of the assignments for overall legibility of letters and overall legibility of words. In addition, the teachers' perceptions about alignment, spacing, and size were obtained using the same scale. The additional criteria were included based on their inclusion on the ETCH-M as non-scored observations.

As would be typical during a referral for occupational therapy evaluation, the teachers were not blinded to which of their students' handwriting they were scoring. Using the same scale, all of the teachers also graded a set of 10 handwriting samples provided by the researchers for purposes of determining inter-rater agreement between the teachers.

### **Data Analyses**

All data were analyzed using SPSS version 24.0 (International Business Machines Corporation [IBM], 2016). Descriptive statistics and graphs were used to describe the initial results. The ETCH-M data were left-skewed, violating assumptions of normality and indicating the need for the use of non-parametric statistical analysis for comparison of distribution between boys and girls. The sample skewness statistic ranged from -0.856 to -2.249 (see Table 1). All of these sample skewness values are statistically significant at the level  $\alpha = 0.05$  to indicate negative skewness in the population. The Mann-Whitney U test was used to analyze the between-group ETCH-M data.

ICCs were computed to determine inter-rater agreement between the teachers. The teachers' perceptions of the students' handwriting were used to divide ratings into two categories: well below to below average for grade level, and average to superior for grade level; it was assumed that students with average or above handwriting would not be referred for occupational therapy services. Receiver operator characteristic curves (ROC) were generated and analyzed to provide insight into choosing cutoffs based on the ETCH-M scores. Multivariate statistical analysis was done using best subset regression analysis (BSRA) and principal component analysis (PCA) to determine if all the subtest scores were necessary for identifying students with deficient handwriting.

### **Results**

The study included 74 participants; 45.9% ( $n = 34$ ) were male and 54.1% ( $n = 40$ ) were female. The participants were predominantly English speaking (94.6%,  $n = 70$ ). The mean age of the participants was 7.6 years with a range of 7.08 to 8.92 years. Most of the participants were right-handed (93.2%,  $n = 69$ ).

Eleven of the participants (14.9%) were reported to have at least one of the following disability diagnoses: ADHD (9.5%,  $n = 7$ ), autism (4.1%,  $n = 3$ ), developmental delay (1.4%,  $n = 1$ ), learning disability (1.4%,  $n = 1$ ), emotional behavioral disability (1.4%,  $n = 1$ ), and other (2.7%,  $n = 2$ ); three participants had two co-morbid diagnoses. Thirteen of the participants were reported to receive resource room help (12.2%,  $n = 9$ ) or special education services (5.4%,  $n = 4$ ); of these, two (2.7%) were reported to receive both resource and special education services. Four students who received resource room help had no reported diagnoses. Eight of the participants were reported to receive related services: speech therapy only (9.5%,  $n = 6$ ), physical therapy and speech therapy (1.4%,  $n = 1$ ), and occupational therapy and speech therapy (1.4%,  $n = 1$ ). Four students receiving speech therapy had no reported diagnosis. The participants could report as many diagnoses, support services, or related services as applied to their individual situation.

Legibility scores were analyzed and are listed in Table 1 in terms of means, standard deviations, medians, and lower quartiles. Speed scores were not reported. The total word mean scores were 88.81 ( $SD = 11.13$ ). Total letter mean scores were 84.30 ( $SD = 10.15$ ). Total numeral mean scores were 89.25 ( $SD = 9.41$ ). The distribution of the scores was left-skewed. We report the mean and standard deviation for purposes of comparison with other results in the literature but note that the median is higher than the mean for all but two subtests because of skewness. For three of four subtests related to word legibility, medians were near 100%, but means were markedly lower because of a small number of low scores.

**Table 1**

*ETCH-M Legibility Subtest Means, Standard Deviations, Mean Ranks, Medians, and Skewness of Scores for All Second-Grade Participants*

ETCH Subtest	M (SD)	Q1	Mdn	Range	Skewness	Std. Error
Lower Case	83.73 (12.37)	80.76	88.40	30.8 – 100	-1.525	.279
Upper Case	81.25 (16.50)	73.09	88.46	19.23 – 100	-1.515	.279
Numeral	90.08 (9.81)	83.33	91.70	58.4 – 100	-1.013	.279
Near Point Word	91.62 (16.22)	80.00	100.00	20 – 100	-2.249	.279
Near Point Letter	89.55 (9.96)	87.45	88.90	55.55 – 100	-1.561	.279
Far Point Word	90.81 (16.94)	80.00	100.00	20 – 100	-2.007	.279
Far Point Letter	85.25 (13.04)	77.77	88.90	44.4 – 100	-.999	.279
Dictation Word	77.03 (28.62)	66.70	83.35	0 – 100	-1.322	.279
Dictation Letter	82.67 (15.88)	73.32	86.70	6 – 100	-1.882	.279
Sentence Word	90.26 (15.81)	85.12	100.00	33.4 – 100	-1.830	.279
Sentence Letter	85.42 (12.64)	80.29	87.75	47.8 – 100	-1.309	.279
Total Word	88.81 (11.13)	83.09	89.50	52.4 – 100	-1.304	.279
Total Letter	84.30 (10.15)	79.02	86.19	54.92 – 99.2	-.856	.279
Total Numeral	89.25 (9.41)	86.75	88.23	52.9 - 100	-1.575	.279

When comparing scores based on gender (see Table 2), there were statistically significant differences across most subtests and total scores with *p* values given in Table 3. Error plot analysis with a 95% confidence interval showed the boys scored consistently lower than the girls (see Figure 1). There were inconsistent differences between the medians and means between the boys and the girls. Medians of all four of the subtests related to word legibility for girls were at 100% with significantly lower means. Medians of two of four of the subtests related to word legibility were at 100% for the boys with significantly lower means.

**Table 2**

*ETCH-M Subtest Means, Standard Deviations, Mean Ranks, and Medians per Gender*

ETCH Subtest	Male N=34			Female N=40		
	M (SD)	Q1	Mdn	M (SD)	Q1	Mdn
Lower Case	80.53 (11.90)	69.20	84.60	86.44 (12.25)	84.60	88.50
Upper Case	78.86 (13.78)	69.86	80.80	83.27 (18.43)	77.88	88.50
Numeral	87.48 (10.70)	75.00	91.66	92.29 (8.51)	91.66	91.70
Near Point Word	88.82 (17.88)	80.00	100.00	94.00 (14.46)	100.0	100.00
Near Point Letter	87.25 (11.44)	83.30	88.90	91.50 (8.15)	88.85	94.40



Far Point Word	86.47 (18.23)	75.00	100.00	94.50 (15.01)	100.00	100.00
Far Point Letter	79.34 (12.45)	70.82	83.30	90.27 (11.44)	88.88	94.40
Dictation Word	68.63 (31.72)	66.70	66.70	84.17 (23.84)	66.70	100.00
Dictation Letter	77.45 (14.20)	66.70	80.00	87.11 (16.04)	81.66	93.30
Sentence Word	86.96 (15.22)	80.00	88.19	93.07 (15.94)	93.17	100.00
Sentence Letter	80.18 (11.85)	73.82	83.70	89.88 (11.65)	85.12	92.60
Total Word	84.82 (10.55)	78.62	88.21	92.21 (10.58)	88.20	94.70
Total Letter	80.26 (9.79)	74.46	81.16	87.73 (9.25)	82.17	90.15
Total Numeral	86.32 (8.40)	82.35	88.20	91.74 (9.59)	88.20	94.10

**Table 3**  
*ETCH Subtest Mann-Whitney U Results of Gender Comparisons*

ETCH Subtest	Mann-Whitney Male vs. Female	
	U	p
Lower Case	6.16	.013*
Upper Case	5.128	.024*
Numeral	4.107	.043*
Near Point Word	3.252	.071
Near Point Letter	4.095	.043*
Far Point Word	5.509	.019*
Far Point Letter	17.308	.000*
Dictation Word	6.102	.014*
Dictation Letter	12.21	.000*
Sentence Word	5.835	.016*
Sentence Letter	16.736	.000*
Total Word	10.676	.001*
Total Letter	12.199	.000*
Total Numeral	11.668	.001*

Note. \*significant at the  $p < .05$  level.

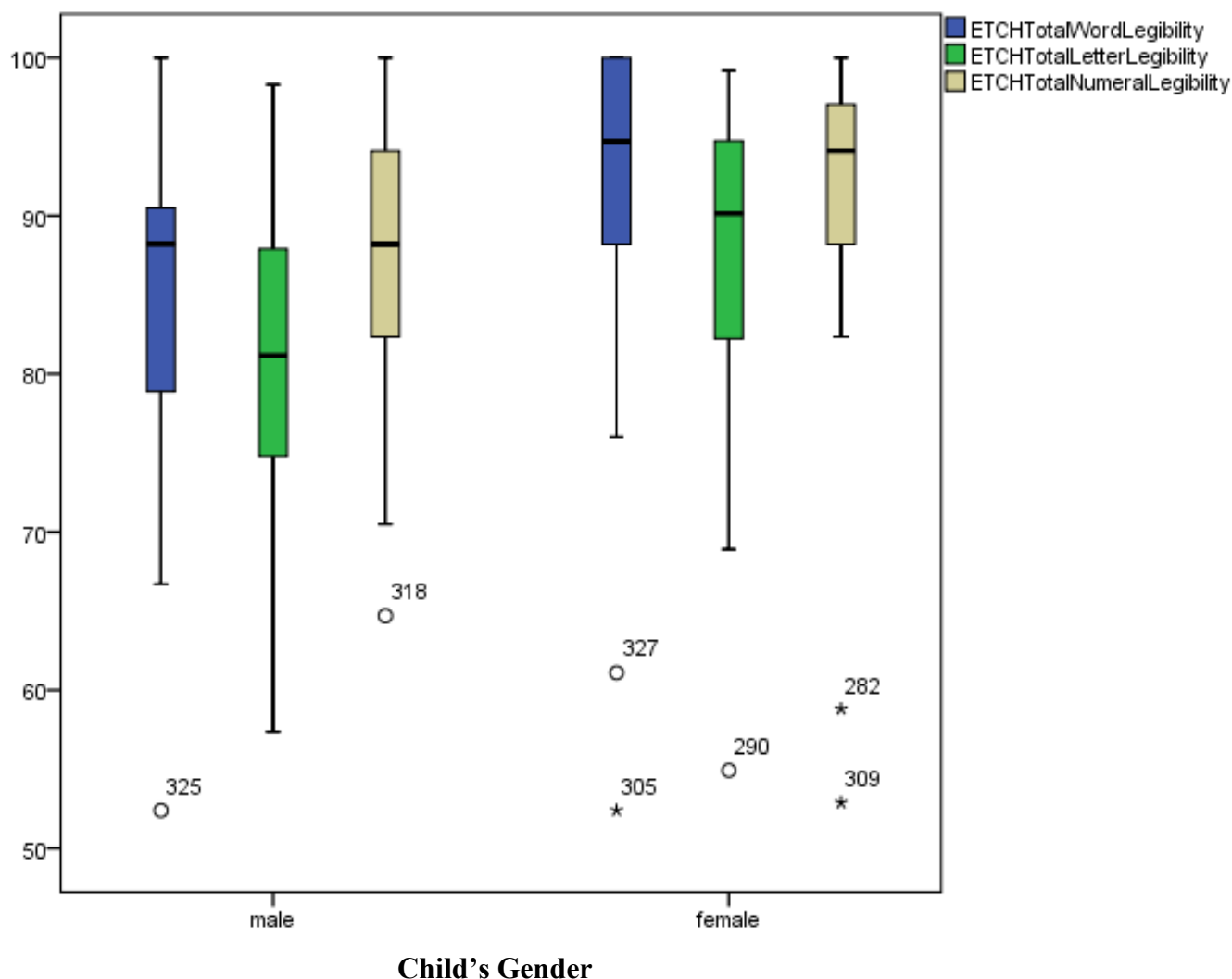


Figure 1. Boxplots of boys compared to girls on total word, letter, and numeral scores.

ICCs were calculated on the teacher scoring of the 10 researcher-provided handwriting samples and revealed a high degree of consistency between the teachers. ICCs for the teachers were .927 for overall legibility, .897 for general legibility of words, and .887 for general legibility of letters. ROCs were generated using the teachers' perceptions of general legibility of words, general legibility of letters, spacing, size, and alignment. The ROCs were analyzed to provide insight into choosing cutoff values based on the ETCH-M scores to best identify those students who would be classified by teachers as well below-age expectancy (see Figures 2 and 3). ROC curves can only provide a rough guideline for a cutoff score and "the final choice of a cutoff, however, would be based on the impact of an incorrect identification" (Portney & Watkins, 2009, p. 639). We first considered the teachers' perceptions of legibility in relation to the ETCH-M total legibility scores, since this was the most direct relationship. We then considered the teachers' perceptions of spacing, size, and alignment in relation to the ETCH-M total legibility scores to supplement our initial findings.

The crude area under the curve for the ETCH-M total letter legibility in relation to the teachers' perceptions of letter legibility was .825 (95% CI = .670, .980). Based on an examination of the data, a cautious recommendation for a cutoff score for total letter legibility based on the teachers' perceptions

of general legibility of letters is 77% (see Figure 2). This choice for a cutoff score yields a sensitivity of 87% and specificity of 73%. Sensitivity measures the probability of a false negative while specificity measure the probability of a false positive. A high sensitivity would yield a low probability of a false negative. Using this cutoff, 87% of the students perceived by teachers as having poor handwriting would be identified by ETCH-M scores of 77% or less. However, 27% of the students with low ETCH-M scores would not be identified by teachers as having poor handwriting.

This choice of cutoff for letter legibility was supported by analysis of the teachers' perceptions of spacing, size, and alignment in relation to the ETCH-M total letter legibility scores. The crude area under the curve for the teachers' perceptions of spacing was .702 (95% CI = .547, .856), for the teachers' perceptions of size it was .760 (95% CI = .619, .900), and for the teachers' perceptions of alignment it was .777 (95% CI = .635, .920). All three yielded potential cutoff scores similar to the total letter analysis. A lower cutoff for letter legibility should be used if increased specificity is desired.

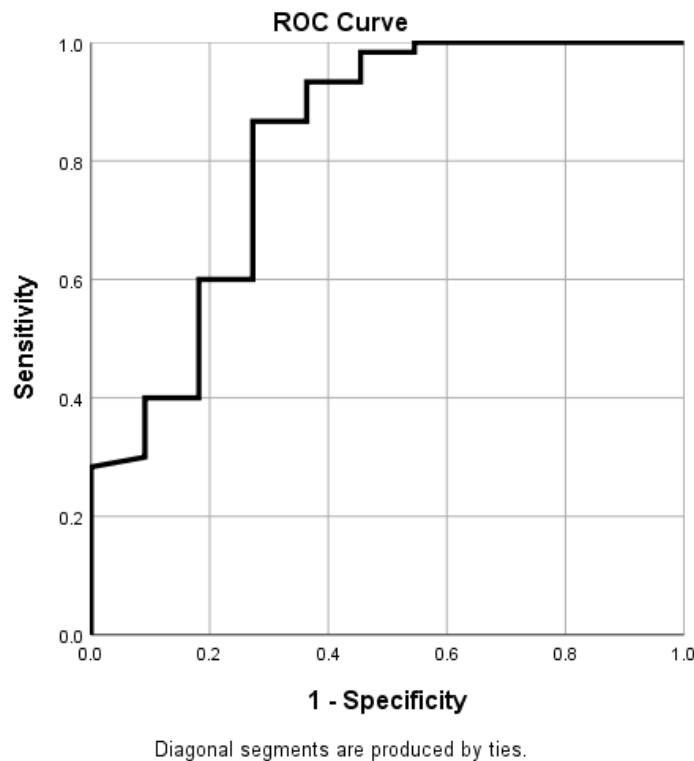


Figure 2. Teachers' perceptions of letter legibility in relation to the ETCH-M total letter legibility.

The teachers' perceptions of overall word legibility was compared to the ETCH-M total word legibility. The crude area under the curve for word legibility was only .682 (95% CI = (.534, .829)). The low value for the area under the curve suggests a low ability to discriminate between those at risk and those not at risk based on this test. A cautious recommendation for a cutoff score for word legibility based on the teachers' perceptions of general legibility of word is approximately 82% (see Figure 3). This choice for a cutoff score yields a sensitivity of 83% and specificity of 43%. Using this cutoff, 83% of the students perceived by teachers as having poor handwriting would be identified by ETCH-M scores of 82% or less. However, 57% of students with low ETCH-M scores would not be identified by teachers as having poor handwriting.

The analysis of the teachers' perceptions of spacing [area = .641, (95% CI = .473, .810)], size [area = .577, (95% CI = .397, .757)], and alignment [area = .618, 95% CI = .431, .810)] in relation to the ETCH- M total word legibility yielded values below .5 in the 95% confidence interval. These criteria were not statistically significant for determining a cutoff value leaving our determination of a potential cutoff score based wholly on the teachers' perceptions of general legibility of words. A lower cutoff for word legibility should be used if increased specificity is desired. A larger dataset is necessary to determine the optimal cutoff.

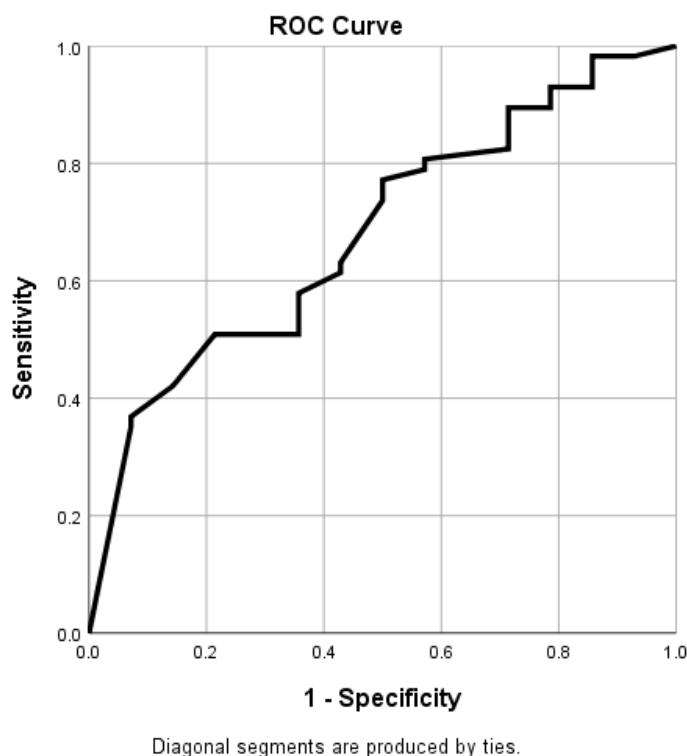


Figure 3. Teachers' perceptions of word legibility in relation to the ETCH-M total word legibility.

Finally, multivariate statistical analysis was done using multiple regression and PCA to investigate whether using a smaller number of subtest scores from the ETCH-M test had the potential to simplify the testing and analysis. PCA is used to investigate whether using a smaller number or different groupings of subtests from the ETCH-M had the potential to simplify the testing and analysis. PCA did not yield helpful dimension reduction.

Best subset regression analysis (BRSA) is used to identify predictor variables. In this case, BRSA was performed to determine which of the ETCH-M subscores were most valuable in predicting the teachers' perceptions of overall legibility. The use of the three total scores (word, letter, and numeral) provided moderate prediction ( $R^2 = 0.487$ ); use of the total letter score alone provided nearly the same predictive value of the three total scores ( $R^2 = 0.457$ ). Of the separate subscores used to compute the total letter score, upper case legibility was the single best indicator ( $R^2 = 0.384$ ).

### Discussion

The results suggest that students in the second grade in the first half of the academic year achieve mean scores on the ETCH-M of high 80 percentages for total legibility of words and numerals, and in

the mid 80 percentages for total legibility of letters. As expected, the mean scores were higher than those reported elsewhere for first-grade students (Feder et al., 2007) and for the mean legibility scores for students in Grades 2 and 3 who had ADHD (Brossard-Racine et al., 2012). The study offers cautious support for the premise that the ETCH-M discriminates adequately between students in different grade levels in comparison to previous reports and that typical students perform better than those with ADHD.

As in other studies (Feder et al., 2007; Graham et al., 1998), the boys did more poorly on the legibility tasks than the girls. Clinically, the difference in average scores between boys and girls raises questions about referral for intervention and whether a different standard should be applied to students based on gender. The boys' median scores for word legibility, though lower than the girls, were six percentage points above the cutoff for well below to below average as perceived by the teachers. The boys' median scores for letter legibility, also lower than the girls, were eight percentage points higher than the suggested cutoff score. The average boys' scores were also above the cutoff scores reported by Brossard-Racine et al. (2012) for word and letter legibility for children with ADHD. It appears that clinicians might expect less legibility from boys, but sufficient discrimination between average and deficit legibility exists.

Cutoff values reported by Brossard-Racine et al. (2012) were similar to our findings for total letter legibility. Brossard-Racine et al. suggested a lower cutoff for word legibility than found in our study. The ROC based on word legibility in our study did not lend itself to a clear cutoff value and a lower value might be a better option if avoiding a false negative is desired.

The analysis using BRSA suggests that the ETCH-M total letter legibility score alone yields nearly the same predictive value of the teachers' perceptions of overall legibility as the three total scores combined. Since all of the ETCH-M subtests contribute to the total letter legibility score, this finding does not help in terms of reducing the number of subtests needed during administration but could simplify interpretation of the ETCH-M results. The findings warrant further exploration to determine if letter legibility alone is sufficient for making determinations about intervention.

### **Limitations**

It should be noted that this convenience sample yielded a participant pool that was in a small geographic location. A larger sample with a wider diversity of participants would increase the confidence of the findings.

The teachers were not blinded to which of their students' handwriting samples they were scoring. The possibility of teacher bias influencing their scoring is recognized. Researchers have shown that factors, such as a student's behavior or ethnicity, can be sources of bias (Bennett, Gottesman, Rock, & Cerullo, 1993; Clark & Zygmunt, 2014; Tenenbaum & Ruck, 2007), as well as the teacher's emotional state when grading student work (Brackett, Floman, Ashton-James, Cherkasskiy, & Salovey, 2013). It is possible that teachers scored individual students better or worse depending on factors unrelated to handwriting. There was reasonably good agreement overall between the teachers' perceptions and the objective ETCH-M scores.

Another potential limitation was that the students were not completing the ETCH-M in their classrooms. As with the pull-out therapies, the students were invited to come with the examiners to another room in the school with which they were familiar. The teachers introduced the examiners to the students and reassured them they would return to their classrooms when finished. The students were told they were not being tested and that they were helping the examiners understand more about second-grade students. Despite attempts to make the situation comfortable for the students, factors such as

leaving the classroom with an unknown adult, using school spaces other than their classrooms, and student sensitivity to testing could have heightened the students' anxiety when completing the ETCH-M and might have influenced the results.

### Implications for Occupational Therapy Practice

The study provides occupational therapists information regarding the performance of typical second-grade students on the ETCH-M. Therapists might expect lower scores from boys than girls. The teachers' perceptions of handwriting skill did discriminate between students who had average to superior handwriting and those who were deemed below average to well below average in relation to the total letter scores, in particular. There was weaker support for a relationship with total word scores. The findings suggest that the total legibility scores on the ETCH-M might correspond to a teacher's perception of legibility. The use of the total letter score alone might be adequate for determining teacher prediction with the ETCH-M scores and might help therapists streamline the evaluation interpretation process. Table 1 is provided to give therapists some information about average performance of second-grade students on the ETCH-M. Future study is needed to see if these results can be generalized to students with different demographic profiles.

---

### References

- Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among development skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology, 85*(3), 478-508. <http://dx.doi.org/10.1037/0022-0663.85.3.478>
- Altemeier, L. E., Abbott, R. D., & Berninger, V. W. (2008). Executive functions for reading and writing in typical literacy development and dyslexia. *Journal of Clinical and Experimental Neuropsychology, 30*(5), 588-606. <https://doi.org/10.1080/13803390701562818>
- Amundson, S. J. (1995). *The Evaluation Tool of Children's Handwriting*. Homer, AK: OT Kids. Retrieved from <http://www.alaska.net/~otkids/ETCH.html>
- American Occupational Therapy Association. (2014). Occupational therapy practice framework: Domain and process (3rd ed.). *American Journal of Occupational Therapy, 68*(Suppl. 1), S1-S48. <http://dx.doi.org/10.5014/ajot.2014.682006>
- Bennett, R. E., Gottesman, R. L., Rock, D. A., & Cerullo, F. (1993). Influence of behavior perceptions and gender on teachers' judgments of students' academic skill. *Journal of Educational Psychology, 85*(2), 347-356. <http://dx.doi.org/10.1037/0022-0663.85.2.347>
- Brackett, M. A., Floman, J. L., Ashton-James, C., Cherkasskiy, L., & Salovey, P. (2013). The influence of teacher emotion on grading practices: A preliminary look at the evaluation of student handwriting. *Teachers and Teaching: Theory and Practice, 19*(6), 634-646. <https://doi.org/10.1080/13540602.2013.827453>
- Brossard-Racine, M., Mazer, B., Julien, M., & Majnemer, A. (2012). Validating the use of the Evaluation Tool of Children's Handwriting – Manuscript to identify handwriting difficulties and detect change in school-age children. *American Journal of Occupational Therapy, 66*(4), 414-421. <https://doi.org/10.5014/ajot.2012.003558>
- Cahill, S. M., McGuire, B., Krundick, N. D., & Lee, M. M. (2014). National survey of occupational therapy practitioners' involvement in Response to Intervention. *American Journal of Occupational Therapy, 68*, e234-e240. <http://dx.doi.org/10.5014/ajot.2014.010116>
- Case-Smith, J., Holland, T., & Bishop, B. (2011). Effectiveness of an integrated handwriting program for first-grader students: A pilot study. *American Journal of Occupational Therapy, 65*, 670-678. <https://doi.org/10.5014/ajot.2011.000984>
- Case-Smith, J., Holland, T., Lane, A., & White, S. (2012). Effect of a co-teaching handwriting program for first graders: One-group pretest-posttest design. *American Journal of Occupational Therapy, 66*, 396-405. <http://dx.doi.org/10.5014/ajot.2012.004333>
- Clark, G. F., Brouwer, A., Schmidt, C., & Alexander, M. (2008). Response to Intervention (RtI) model: Using the Print Tool to develop a collaborative plan. *OT Practice, 52*(4), 248-255.
- Clark, P., & Zygmunt, E. (2014). A close encounter with personal bias: Pedagogical implications for teacher education. *Journal of Negro Education, 83*(2), 157-161. <https://doi.org/10.7709/jnegroeducation.83.2.0147>
- Diekema, S. M., Deitz, J., & Amundson, S. J. (1998). Test-retest reliability of the Evaluation Tool of Children's Handwriting – Manuscript. *American*

- Journal of Occupational Therapy*, 52(4), 245-55.  
<https://doi.org/10.5014/ajot.52.4.248>
- Donica, D. K. (2015). Handwriting without Tears®: General education effectiveness through a consultative approach. *American Journal of Occupational Therapy*, 69, 6906180050.  
<https://doi.org/10.5014/ajot.2015.018366>
- Donica, D. K., Larson, M. H., & Zinn, A. A. (2012). Survey of handwriting instruction practices of elementary teachers and educational programs: Implications for occupational therapy. *Occupational Therapy in Health Care*, 26(2-3), 120-137.  
<https://doi.org/10.3109/07380577.2012.693244>
- Feder, K. P., Majnemer, A., Bourbonnais, D., Blayney, M., & Morin, I. (2007). Handwriting performance on the ETCH-M of students in a grade one regular education program. *Physical & Occupational Therapy in Pediatrics*, 27(2), 43-62. [https://doi.org/10.1080/J006v27n02\\_04](https://doi.org/10.1080/J006v27n02_04)
- Feder, K. P., & Majnemer, A. (2003). Children's handwriting evaluation tools and their psychometric properties. *Physical & Occupational Therapy in Pediatrics*, 23(3), 65-84. [https://doi.org/10.1300/J006v23n03\\_05](https://doi.org/10.1300/J006v23n03_05)
- Feder, K., Majnemer, A., & Synnes, A. (2000). Handwriting: Current trends in occupational therapy practice. *Canadian Journal of Occupational Therapy*, 67, 197-204.  
<https://doi.org/10.1177/000841740006700313>
- Gerde, H. K., Foster, T. D., & Skibbe, L. E. (2014). Beyond the pencil: Expanding the occupational therapists' role in helping young children to develop writing skills (Report). *Open Journal of Occupational Therapy*, 2(1), 1-19.  
<https://doi.org/10.15453/2168-6408.1070>
- Grace-Frederick, L. (1998). *Printing legibility, pencil grasp, and the use of the ETCH-M*. [Unpublished Master's thesis, Boston University].
- Graham, S., Berninger, V., Weintraub, N., & Schafer, W. (1998). Development of handwriting speed and legibility in grades 1-9. *Journal of Educational Research*, 92(1), 42-52.  
<https://doi.org/10.1080/00220679809597574>
- Hammerschmidt, S. L., & Sudsawad, P. (2004). Teacher's survey on problems with handwriting: Referral, evaluation, and outcomes. *American Journal of Occupational Therapy*, 58(2), 185-192.  
<https://doi.org/10.5014/ajot.58.2.185>
- 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, 13-25.  
<https://doi.org/10.2182/cjot.2011.78.1.3>
- International Business Machines Corporation [IBM]. (2016). IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.
- Milone, M. (2007). *Test of Handwriting Skills Manual*. Los Angeles, CA: Western Psychological Services.
- Ohl, A. M., Graze, H., Weber, K., Kenny, S., Salvatore, C., & Wagreich, S. (2013). Effectiveness of a 10-week Tier-1 Response to Intervention program in improving the fine motor and visual-motor skills in general education kindergarten students. *American Journal of Occupational Therapy*, 67, 507-514.  
<http://dx.doi.org/10.5014/ajot.2013.008110>
- Olsen, J. Z., & Knapton, E. F. (2016). *The Print Tool*. Retrieved from <https://shopping.lwtears.com/category/theprinttool>
- Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice* (3rd ed.). Upper Saddle River, NJ: Pearson Education, Inc.
- Reeder, D. L., Arnold, S. H., Jeffries, L. M., & McEwen, I. R. (2011). The role of occupational therapists and physical therapists in elementary school system early intervening services and response to intervention: A case report. *Physical and Occupational Therapy in Pediatrics*, 31(1), 44-57.  
<https://doi.org/10.3109/01942638.2010.497180>
- Reeder, D. L., Arnold, S. H., Jeffries, L. M., & McEwen, I. R. (2011). The role of occupational therapists and physical therapists in elementary school system early intervening services and response to intervention: A case report. *Physical and Occupational Therapy in Pediatrics*, 31(1), 44-57.  
<https://doi.org/10.3109/01942638.2010.497180>
- Reisman, J. (1999). *Minnesota Handwriting Assessment*. San Antonio, TX: Psychological Corporation.
- Rosenblum, S. (2018). Inter-relationship between objective handwriting features and executive control among children with developmental dysgraphia. *PLoS ONE*, 13(4), e0196098.  
<https://doi.org/10.1371/journal.pone.0196098>
- Sudsawad, P., Trombly, C. A., Henderson, A., & Tickle-Degen, L. (2001). The relationship between the Evaluation Tool of Children's Handwriting and teachers' perceptions of handwriting legibility. *American Journal of Occupational Therapy*, 55(5), 518-523.  
<https://doi.org/10.5014/ajot.55.5.518>
- Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99(2), 253-273. <https://doi.org/10.1037/0022-0663.99.2.253>
- Woodward, S., & Swinth, Y. (2002). Multisensory approach to handwriting remediation: Perceptions of school-based occupational therapists. *American Journal of Occupational Therapy*, 56, 305-312.  
<http://dx.doi.org/10.5014/ajot.56.3.305>
- Zylstra, S. E., & Pfeiffer, B. (2016). Effectiveness of a handwriting intervention with at-risk kindergarteners. *American Journal of*

*Occupational Therapy, 70, 7003220020.*

<https://doi.org/10.5014/ajot.2016.01882>