Document Title:	**Development of Individual Handwriting Characteristics in ~1800 Students: Statistical Analysis and Likelihood Ratios that Emerge Over an Extended Period of Time**
Author(s):	**Lisa M. Hanson**
Document No.:	**250165**
Date Received:	**August 2016**
Award Number:	**2010-DN-BX-K212**

DEVELOPMENT OF INDIVIDUAL HANDWRITING CHARACTERISTICS IN ~1800 STUDENTS:

STATISTICAL ANALYSIS AND LIKELIHOOD RATIOS THAT EMERGE OVER AN EXTENDED PERIOD OF TIME

NIJ Fundamental Research to Improve Understanding of the Accuracy, Reliability, and Measurement Validity of Forensic Science Disciplines (FY2010);

MN Bureau of Criminal Apprehension; DUNS—804886729,

Forensic Scientist 3, Lisa M Hanson, May 15, 2016

Purpose of the Project

For decades, questioned document examiners (QDE) have been conducting handwriting comparisons and testified to their conclusions in civil and criminal courts in the US and internationally. The task of handwriting identification is based on an assumption of individuality: No two people write the same way and no one person writes exactly the same way twice [1], [2], [3]. Statistical support for this identifiability in the past using two types of data: (i) random samples from a population representative of the United States (4), and (ii) samples from twins to take the genetic factor into account (5).

Both previous approaches have a bias which we attempt to address in the study reported here. General population across the United States consists of people from different school systems or even different cultures. Therefore, they may learn how to write by different ways of teaching, or they may just learn to write in different writing systems. So there is a large variation between their handwritings. However, the purpose of this study takes a third approach – using writing samples from a population, all who are in the same age groups and instructed in the same manner.

This study focuses on analysis over time, specifically children's handwriting as they develop their writing skills during early grades. Because the task of writer identification is based on the assumption of individuality of handwriting, it is important to be aware of how individual handwriting characteristics are assumed to develop. Students are taught to write in primary school; hand printing begins in kindergarten and cursive writing begins in second grade. They are taught a style of hand writing (e.g., D'Nealian, Zaner-Bloser, etc.) by copying letter and number formations from a copy book or from an instructional banner posted in their classroom. The writing system determines *Class Characteristic* of all the students; in the case the student's *Class Characteristic* (properties common to a group) will be the letter formations instructed using Zaner-Bloser. An example of the writing is shown in Fig. 1. [Figure 1 here].

Over time, with continued instruction and practice, the students' writing skills increase as well as their confidence; they will stop copying the letters and numbers and instead begin writing from memory. It's at this time, when students begin to write from memory, that students are believed to begin developing their *Individual Characteristics* (unique to individual) [6].

Questions this project aims to answer are: (i) at what ages do individual handwriting characteristics begin to develop; (ii) at what rate do these individual handwriting characteristics develop; (iii) what are the most common (less unique) individual characteristics that develop and (iv) what are the least common individual (more unique) individual characteristics that develop. Such a temporal analysis will not only lead to an understanding as to how individual handwriting characteristics develop and come to be habitual, but it will also provide a clearer understanding of the accuracy, reliability and measurement validity of the handwriting comparison processes.

Project Subjects

Handwriting samples were collected from school age children in second grade (~7 years old), third grade (~8 years old) and fourth grade (~9 years old) because writing instructions (lessons) begin and then continue to be practiced during these three grades. Hand printed and cursive writing samples were collected from all the students, approximately 1,800 subjects, as they just began to learn (2nd graders) or had just recently learned (3rd and 4th graders) to write using cursive writing. The first year's writing collection was gathered during spring of the 2011-2012, the second year during spring of the 2012-2013 and the third year during the spring of 2013-2014. The Minnesota Independent School Districts (MNISD) participating in this study were chosen because they had low "family move in/move out" ratios in their districts over the past 10 years. Both school districts chosen were teaching their students to write using the Zaner-Bloser method of writing; the most common writing method being taught throughout the United States at the present. Writing samples were collected each spring in order to obtain consistent data sampling, necessary to document minute changes in hand printing and cursive writing skills and habits, as the students' abilities change over time. The students were asked to produce the sample paragraph a total of four times (two times hand printed; two times cursive). Most students achieved this task; some wrote less, some wrote more.

Project Design

The PI, in conjunction with MNISD #832 and #833, University Professor, Dr. Sargur N. Srihari and Dr. Greg Ball gathered all the handwriting and data analysis by completing the following:

1. The PI and the ISDs worked together assigned each student a unique participant identification number used throughout the study.
2. Each spring handwriting and hand printing samples were gathered from a large number of students as they were learning or had just learned how to write. These same students were followed through all three years of their primary education writing career, as were available.
3. Hand writing forms were produced by the PI in order to control the document writing area and line spacing variables. Writing forms for second and third graders were produced with a solid baseline and top line and a dashed middle line, similar to instruction forms used during class instructions. Third grade and above all used single lined forms.
4. Each student was asked to produce 4 copies of the same paragraphs: 2 copies in cursive writing and 2 copies in hand printing.
5. **First year**: PI gathered only hand printing writing from students in 2nd graders and hand printing and cursive writing from 3rd and 4th graders; **Second year**: PI gathered hand printing and cursive

writing from 3^{rd}, 4^{th} and 5^{th} graders; **Third year**: PI gathered hand printing and cursive writing from 4^{th}, 5^{th} and 6^{th} graders

6. All writing samples were scanned as whole page documents, digitized and the all of the words ***and*** were extracted and saved as separate images. If less than 5 ***and*** were present the sampling was not analyzed for that student that year.
7. Each student's ***and*** was analyzed using two different approaches. In the first approach, the distribution of the characteristics of the letters in the word ***and*** was studied using characteristic values, assigned by human board certified questioned document examiners, using a truthing tool. This resulting data was analyzed first by constructing probabilistic graphical models (7), from which information theoretic measures were computed to determine the range of variations (8,9). In the second approach, the entire sample of writing, rather than a single word, was examined by the use of an automated system for handwriting comparison.

Methods

The students write a paragraph four times each year; two times in cursive writing and two times by hand printed. The paragraph consists of the following text:

The brown fox went into the barn where he saw the black dog. After a second, the black dog saw the fox too. The brown fox was fast and quick. The black dog was not fast and he lost the fox. The fox hid in a hole and waited for the black dog to go home. After the black dog went home, the fox was able to go to the hole he called home and saw all the other foxes. The other foxes were glad to see him and they all asked him to tell them about his day.

The word ***and*** appears five times in the paragraph. This data gathering process produced an average of 5 to 10 samples of the word per student in hand printed and in cursive (except for second graders) to be analyzed using a truthing tool, which assigns numbers to the sub-categories involved with the 12 individual handwriting characteristics (7). The word *and* is one of few words that children write over and over during their early writings and have therefore began to develop individual handwriting characteristics rather than words that are not written often.

These QDEs assigned different characteristics to handwriting manually, depending on whether the writing is cursive or hand printed. The Zaner-Bloser Copy Book Handwriting Style (10), (see Fig. 2), is the hand printing / hand writing copy book style being taught in elementary schools in Minnesota to introduce children to writing. A handwritten letter, or a combination of letters such as the word *and*, can be represented by a set of D characteristics, $\mathbf{X} = \{X_i\}$, i =1,..,D where characteristic X_i takes one of d_i discrete values. The characteristics assigned to our dataset of children's handwriting is represented in Fig. 3. [Figure 2 here][Figure 3(a-b) here]

Samples of handwritten *and* by children of grade 2, 3 and 4 and their respective feature values are shown in Fig. 4. [Figures 4 here]

The data sets gathered using the truthing tool were then forwarded to Dr. Srihari and Dr. Ball for statistical analysis. Two types of analyses were conducted: one based on similarity between samples, and another based on probability distribution of the samples. The distribution of the characteristics was analyzed first by constructing probabilistic graphical models of the data. Information theoretic measures, entropy and relative entropy, were computed using samples generated from the models. The measures are used to document the changes that take place in each student's own writing from year to year, as well as the changes that occur between the time the students learn (copy) from the copy book and then stray away as time passes between grades.

Data Analysis

In studying the differences between populations of different grades, a statistical measure, *relative entropy (K-L divergence),* becomes useful. It measures the change in disorder between two populations. Given a vector **x** of discrete characteristics, and distributions $p(\mathbf{x})$ and $q(\mathbf{x})$, relative entropy is represented as:

$$KL(p||q) = \sum_{\mathbf{x}} p(\mathbf{x})(\ln p(\mathbf{x}) - \ln q(\mathbf{x}))$$

Where ln represents natural logarithm and the summation is over all $\prod_{i=1}^{D} d_i$ values of **x**, $KL(p||q)$ represents information lost in nats (or in bits if logarithm has base 2) when q is used to represent p.

Since we do not have the distributions but have samples $\{\mathbf{x}_1,.., \mathbf{x}_N\}$, the sampling version of relative entropy is:

$$KL(p||q) = \frac{1}{N}\sum_{i=1}^{N}(\ln p(\mathbf{x}_i) - \ln q(\mathbf{x}_i))$$

The sampling version still needs distributions for computing K-L divergence. For this we constructed Bayesian Networks for each of the two distributions by using a causal Bayesian network structure learning algorithm (9).

The relative entropies between different grades for hand printed and cursive handwriting data sets are shown in Table 1. Here it is seen that the relative entropy between Grade 2 (G_2) to Grade 3 (G_3) is smaller than the relative entropy between Grade 2 (G_2) and Grade 4 (G_4). This reflects the development of individuality as the grades progress. [Table 1 here]

Conclusion for Approach 1

The extent of change between grades is measured by relative entropy: the relative entropy of hand printed between grades 2 and 3 is smaller than the relative entropy between grades 3 and 4, and grades 2 and 4.

A somewhat similar effect is seen with the mean of each distribution: (i) the mean for hand printed is closer to the ideal (Zaner-Blozer) in Grade 3 compared to Grade 2, but moves further away for Grade 4. However the mean for cursive is closer to the ideal (Zaner-Bloser) for Grade 4 compared to Grade 3-- which may not be as significant as the increase in entropy.

Approach 2

The automated analysis setup is described by the flowchart in Fig. 5. This analysis focused on hand printed and cursive samples from children in grade 2 during 2011-12, and from the same children in grades 3 and 4 in the years 2012-13 and 2013-14. Table 2 shows the number of samples used. The schools are coded as follows: A= Liberty Ridge, B= Middleton, C= Red Rock, and D= Wildwood. N_i^h and N_i^c are the numbers of hand printed and cursive writing samples in grade i. [Figure 5 here] [Table 2 here]

Handwriting samples were scanned as grayscale images in PNG format. There are three types of noise: printed instructions at the top, black vertical line in the left most part, and the dashed lines and solid lines crossing the letters, whose presence can influence accuracy of automatic comparison. So pre-processing is needed before comparison as shown in Fig. 6(a), which is the resulting image of Fig. 1(a).

Firstly, the source image was cropped to eliminate instructions and vertical lines. After converting the input to a black and white reversed binary image, dashed lines can be easily removed by comparing the length with a threshold. Since dashed lines are built up by small connected parts, they can be removed while preserving handwriting strokes. Solid lines are removed with the help of a mask to preserve the intersection parts between the lines and handwritings. After multiplied by a predetermined weaken rate, the binary image is compressed vertically to build a mask, in which the intersection parts are intensified. Then the means of each row are deducted from the pixel values. Therefore, when the mask is stretched back to its original size, the intersection parts' information can still be preserved. So when the weakened binary image is multiplied by the mask, and converted to binary image again using a restoration threshold, the solid lines can be removed while keeping strokes integral.

This method can get good results if the images are scanned in upright direction. However, the lines in the scanned images are not perfectly horizontal, instead have a tiny slope. Thus, a line removal parameter, *number of batches*, is introduced. The column vectors of the binary image is first separated evenly into a certain number of batches, so that lines can be regarded as horizontal in each batch. Then the batches are processed one by one to remove the lines (11).

The parameter, #batch can influence the degree of line-removal. Fig. 6(b-e) illustrates the results with different values of #batch. With #batch= 10, some relatively long segments near the edge are not removed. With #batch= 20, there are only few noticeable line segments left. When #batch= 40, very few short lines can be found near the edge. With #batch= 80, solid lines and dash lines crossing the words are completely removed, but some other lines are removed incorrectly. This is because when batch number is small, we separate a relatively large number of columns into one batch, so we may mistakenly determine a short line as strokes. While when batch number is large, relatively small number of columns are assigned to one batch, so some long horizontal strokes, like the horizontal line of letter t, may be regarded as lines instead of strokes. It was concluded that removing less leads to segments still left. Removing too much leads to lines being removed incorrectly, e.g., horizontal line of letter t, and continuous strokes becoming disconnected. So, we chose 20 or 40 for 8-bit images, and 10 for 16-bit images. Additionally, line removal was improved in the conversion to binary image part considering that the source images were in two formats: 8-bit and 16-bit. [Figure 6 here]

Log Likelihood Ratio Calculation

For performing comparisons of writing samples we used the CEDAR-FOX software system (5). The indexed input images were firstly mapped with a transcript, containing the paragraph which children were asked to copy. Then based on the recognized characteristics, the system computes three types of features vectors, macro-features, micro-features, and style features, to determine the strength of whether the two inputs are written by the same writer or not, which is measured by log-likelihood ratio (LLR). Examples of macro-features are gray-scale entropy, slant, and height. Micro-features represent handwriting characteristics such as stroke and structure, and style features come from pairs of letters, known as *bigrams*.

The dissimilarity of the two input images is measured by the distance between their feature vectors, which is computed by "Correlation" measure:

$$D(X,Y) = \frac{1}{2} - \frac{S_{11}S_{00} - S_{10}S_{01}}{2\sqrt{(S_{10}+S_{11})(S_{01}+S_{00})(S_{11}+S_{01})(S_{00}+S_{10})}}$$

where $S_{ij}(i,j \in \{0.1\})$ represents the number of matches of i in vector X and j in Y. The distance data is represented by Gaussian and Gamma distributions, and parameters are estimated by maximum likelihood estimation method and stored in the software system. Then the likelihood of same writer L_s, and the likelihood of different writer L_d can be expressed as:

$$L_s = \prod_{i=1}^{c} \prod_{j} \prod_{k} p_s(d_i(j,k))$$

$$L_d = \prod_{i=1}^{c} \prod_{j} \prod_{k} p_d(d_i(j,k))$$

where c is the number of writing elements, $d_i(j,k)$ is the distance between the jth element e_i appearing in the first document and the kth e_i in the second, and p_s and p_d are the probability density functions of distances for same writer and different writer respectively. Then the log-likelihood ratio is given by the software as:

$$LLR = \sum_{i=1}^{c} \sum_{j} \sum_{k} [\ln p_s(d_i(j,k)) - \ln p_d(d_i(j,k))]$$

Therefore, a positive LLR means the likelihood ratio is larger than 1, so it is indicative of the same writer; while a negative LLR refers that the likelihood ratio is less than 1, which means the two inputs are from different writers. When LLR is close to zero, it means the answer is ambiguous (5).

Results

Two types of comparisons were made: (ii) samples collected in two consecutive years by the same children, and (ii) samples from different children in the same grade. The numbers of comparisons made were $N_{i,j}^h$ for hand printed samples and $N_{i,j}^c$ for cursive samples, where $N_{i,j}^h = \min\{N_i^h, N_j^h\}$, and $N_{i,j}^c = \min\{N_i^c, N_j^c\}$. A total of 2755 comparisons were made: 1974 of them for hand printed writing, and 781 for cursive writing. The mean and standard deviation of the computed LLRs are shown as Tables 3-- 4, where Table 3 is the results of comparing same child in different grades and Table 4 is for comparing different children in the same grade. [Table 3 here] [Table 4 here]

Fig. 7 graphically illustrates the means and standard deviations for each comparison type: means are represented as crosses and standard deviations as vertical lines. Fig. 7(a) is from comparing the same children's hand printed writing in two consecutive years: the vertical lines are for comparisons between grade 2 and grade 3, grade 3 and grade 4, and grade 2 and grade 4. Fig. 7(b) also refers to the comparisons between the same children but wrote in cursive style. Fig. 7(c) and Fig. 7(d) illustrate the results for comparing two images from two different children in the same grade. For hand printed, results for grade 2, grade 3, and grade 4 are shown in Fig. 7(c). While for cursive, Fig. 7(d) shows the results from children in grade 3 and the results when they moved up to grade 4. [Figure 7 here]

An example of handwriting for the same child MAA2022VA in different years is shown in Fig. 8: (a) is from grade 2 in 2011-2012, (b) is from the same child in grade 3 in 2012-2013 (c) is from the third year. Comparing images from grades 2 and 3, LLR=51.71, and between grades 3 and 4, LLR= -105.15. Another example of comparisons between different children in the same grade is illustrated in Fig. 9. The three pairs of images are from the same two children, HUA2022CH and IAC2022SA, in grades 2-4 respectively. Comparing each pair, we get low LLR values of -18.64, -101.99, and -178.74, which are smaller with each higher grade. [Figure 8 here] [Figure 9 here]

To determine statistical significance, we use the chi-squared goodness of fit test, to compare different distributions (12). The data are divided into bins, which have at least 5 elements in each of them. Then the test statistic is calculated as

$$X^2 = \sum_{i=1}^{k} (O_i - E_i)^2 / E_i$$

where k is the number of bins, O_i and E_i are the frequency count for the i^{th} bin of the two distributions. The test statistic follows a chi-square distribution. Therefore, the null hypothesis that the two distributions are consistent is rejected if the calculated test statistic X^2 is larger than the chi-square critical value, which can be determined in a chi-square distribution table with $k - 1$ degrees of freedom and significance level α.

We used the chi-squared goodness of fit test for three kinds of comparisons of LLR values: (i) between different grades' hand printed results, (ii) between different grades' cursive results, and (iii) between the two kinds of writing styles in the same grade. Through the test, we get the test statistic for the null hypothesis $H_0 =$ same distribution, as shown in Table 5. According to the chi-square distribution table with α=0.05, null hypothesis of same distribution is rejected. Thus we conclude that all of the compared result distributions are statistically different. [Table 5 here]

The results for different schools are also tabulated. Comparison of the same child's handwriting in two consecutive years is shown in Tables 3. The average LLR between grades 2 and 3 for hand printed writing is -15.85 showing some change but between grades 3 and 4 it is -115.05 indicating a much more marked difference. As for cursive writing between grade 3 and grade 4, the average LLR value is -52.85. LLR values when different grade 2 children are compared are summarized in Table 4. The mean of -57.69

is much larger than the mean of -15.85 in Table 3 for grade pair 2 and 3, implying that different children in the same grade write more differently from each other when compared to the same child in different grades.

Tables 4 also summarizes comparisons between pairs of different children in grades 3 and 4. The mean values of -87.17 and -245.16 for hand printed writing, and -45.12 and -162.98 for cursive writing, indicate that children write more differently as they grow up. Consistently, in Fig. 7(a), the cross on the second vertical line is lower than the one on its left. This shows that when comparing hand printed writings from the same child in different years, the mean LLR of grades 3 to 4 has a larger absolute value than the mean LLR of grades 2 to 3. This also indicates children change their handwriting to a larger degree from grade 3 to 4 than from grade 2 to 3. The crosses in Fig. 7(c) and Fig. 7(d) refer to the means for different children. Observing the three crosses in Fig. 7(c), which are for hand printed writings in grades 2, 3 and 4, or the two in Fig. 7(d), which are for cursive writings in grade 3 and 4, we see that both are progressively lower.

The standard deviations, which represent the amount of variation of the data from its average, are also interesting to observe. They are tabulated as well as indicated by the lengths of the vertical lines in Fig. 7. For the comparisons of the same child, the degree of dispersion is more when children move from grade 3 to 4 than from grade 2 to 3. For the comparisons between different children, dispersion is more for higher grades. This again is consistent with development of handwriting individuality.

Conclusion for Approach 2

We compared both hand printed and cursive handwriting samples of children progressing from grades 2 to 4 to determine the degree of individuality as children develop. We considered the entire writing sample image as input yielding an LLR value for each comparison. Since the quality of images plays an important role, a line removal algorithm was carefully chosen. Resulting LLR values were analyzed using chi-squared goodness of fit tests. They indicate that as children move to higher grades, they gradually begin to form their own writing styles. This provides a strong justification that handwriting becomes more individualistic with age even when children are taught with the same writing style.

Implications for Criminal Justice Policy and Practice

Forensic handwriting examinations are often an important part of criminal and civil cases. Threatening letters, bomb threats, check fraud, homicides, and controlled substance cases, just to name a few. During the past decade, Daubert and Frye Mack hearings have become a common occurrence across the United States involving forensic handwriting comparisons. The collection of these writing samples, along with the measurements of the individual hand writing characteristics as they developed is the beginning of a true statistical model that may be used to scientifically prove why forensic handwriting comparisons are possible. In addition, it may also be possible in the future to continue data mining by examining new words, located within the previously collected paragraphs. By continuing to gather these measurements and statistics, data gathered can continue to statically prove how individual handwriting characteristics develop and how the combinations of these individual handwriting characteristics develop into individual handwriting styles. This study has made it possible replace "theory" with solid statistics that are scientifically accurate and reliable. The outcome of this research is important for hand writing examiners as well as for all courts and legal areas that involve forensic handwriting examinations.

References

1. A. Osborn, Questioned Documents, Nelson Hall Pub, 1929.

2. O. Hilton. Scientific Examination of Questioned Documents. Revised Edition. CRC Press, 1993.

3. R. A. Huber and A. M. Headrick. Handwriting Identification: Facts and Fundamentals. CRC Press, 1999.

4. S. N. Srihari S. Cha, H. Arora and S. J. Li. Individuality of handwriting. J Forensic Sci, 2002, 47(4): 856-872.

5. S. N. Srihari, C. Huang, and H. Srinivasan. On the discriminability of the handwriting of twins. J Forensic Sci, 2008, 53(2): 430-446.

6. Srihari S N, Singer K. Role of automation in the examination of handwritten items. Pattern Recognition, 2014, 47(3): 1083-1095.

7. M. Puri, S. N. Srihari and L. Hanson. Probabilistic Modeling of Children's Handwriting. In: Proc. Document Recognition and Retrieval XXI, San Francisco, CA, Feb. 2014.

8. S. N. Srihari and Z. Xu. Development of Handwriting Individuality: An Information-Theoretic Study. In: Proc. Int. Conf. Frontiers in Handwriting Recognition, Crete, Greece, 2014: 601-606.

9. Z. Xu and S. N. Srihari. Bayesian network structure learning using causality. In: Proc. Int. Conf. Pattern Recognition, 2014.

10. J. S. Kelly and B. Lindblom. Scientific Examination of Questioned Documents. Second Edition, CRC Press, 2006.

11. J. Chu and S.N. Srihari. Writer identification using a deep neural network. In: Proc. Int. Conf. Computer Vision, Graphics and Image Processing, Bangalore, 2014.

12. NIST/SEMATECH e-Handbook of Statistical Methods. http://www.itl.nist.gov/div898/handbook/.

TABLES

Table 1

Differences (KL values) of “and” characteristics of students in different grade pairs

Grade Pair	Hand printed	Cursive
$KL(G_3\|\|G_2)$	0.32	-
$KL(G_4\|\|G_3)$	0.41	0.68
$KL(G_4\|\|G_2)$	0.64	-

Table 2

Numbers of full-page samples from each year, school, and writing type (hand printed/cursive)

N_i^h = No. of hand printed samples from grade i

N_i^c= No. of cursive samples from grade i

Grade (Year)	School / Type	A	B	C	D	Total
G_2 (2011-2012)	N_2^h	57	87	73	149	366
	N_2^c	0	0	0	0	0
G_3 (2012-2013)	N_3^h	57	87	73	149	366
	N_3^c	40	59	50	114	263
G_4 (2013-2014)	N_4^h	47	73	54	122	296
	N_4^c	40	59	50	114	263

Table 3

Results of comparing same child full page samples in different grades (LLR values)

$G^s_{i,j}$ = Comparison of same child in grades i and j

$N^h_{i,j}$ = No. of comparisons of hand printed samples between grades i, j

$N^c_{i,j}$= No. of comparisons of cursive samples between grades i, j

$\mu^h_{i,j}$ = Mean of LLR from comparing hand printed samples of grades i, j

$\mu^c_{i,j}$ = Mean of LLR from comparing cursive samples of grades i, j

$\sigma^h_{i,j}$ = Std. deviation of LLR from comparing hand printed samples of grades i, j

$\sigma^c_{i,j}$ = Std. deviation of LLR from comparing cursive samples of grades i, j

Grade pair $G^s_{i,j}$	School / LLR	A	B	C	D	Total
$G^s_{2,3}$	$N^h_{2,3}$	57	87	73	149	366
	$\mu^h_{2,3}$	-9.66	-4.08	-2.87	-31.13	-15.85
	$\sigma^h_{2,3}$	95.66	93.50	101.69	75.89	89.42
$G^s_{3,4}$	$N^h_{3,4}$	47	73	54	122	296
	$\mu^h_{3,4}$	-87.69	-109.48	-83.85	-142.50	-115.05
	$\sigma^h_{3,4}$	140.32	113.67	219.60	99.04	140.15
	$N^c_{3,4}$	40	59	50	114	263
	$\mu^c_{3,4}$	-72.89	-59.35	-33.04	-51.39	-52.85
	$\sigma^c_{3,4}$	104.24	45.51	49.64	53.95	62.54
$G^s_{2,4}$	$N^h_{2,4}$	47	73	54	122	296
	$\mu^h_{2,4}$	-161.17	-121.15	-171.50	-153.57	-150.45
	$\sigma^h_{2,4}$	107.17	101.85	136.79	92.07	107.15

Table 4

Results of comparing different children full page samples in same grade (LLR values)
G_i^d = Comparison of different children in grade i
μ_i^h = Mean of LLR from comparing hand printed samples of grades i
μ_i^c = Mean of LLR from comparing cursive samples of grades i
σ_i^h = Std. deviation of LLR from comparing hand printed samples of grades i
σ_i^c = Std. deviation of LLR from comparing cursive samples of grades i

Grade G_i^d	School / LLR	A	B	C	D	Total
G_2^d	N_2^h	56	86	72	148	362
	μ_2^h	-63.31	-42.50	-52.16	-67.08	-57.69
	σ_2^h	90.26	56.51	60.31	51.10	62.24
G_3^d	N_3^h	56	86	72	148	362
	μ_3^h	-75.46	-57.57	-55.47	-124.22	-87.17
	σ_3^h	79.64	83.70	81.84	87.90	89.85
	N_3^c	39	58	49	113	259
	μ_3^c	-68.81	-47.20	-19.30	-47.07	-45.12
	σ_3^c	36.70	34.81	30.00	30.25	35.28
G_4^d	N_4^h	46	72	53	121	292
	μ_4^h	-201.43	-184.35	-217.05	-310.28	-245.16
	σ_4^h	260.73	177.86	241.85	176.15	211.20
	N_4^c	39	58	49	113	259
	μ_4^c	-238.89	-165.74	-95.40	-164.68	-162.98
	σ_4^c	92.76	125.95	142.84	104.19	122.56

Table 5

Results of chi-squared goodness of fit tests for H_0 = same distribution

$G_{i,j}^{sh}$ = Comparison of hand printed of same child for grade i and j

$G_{i,j}^{sc}$ = Comparison of cursive of same child for grade i and j

G_i^{dh} = Comparison of hand printed of different children for grade i

G_i^{dc} = Comparison of cursive of different children for grade i

	X^2	Reject H_0 ?
Between different pairs of comparisons for hand printed		
$G_{2,3}^{sh}$ vs $G_{3,4}^{sh}$	561.42	Yes
$G_{3,4}^{sh}$ vs $G_{2,4}^{sh}$	85.75	Yes
G_2^{dh} vs G_3^{dh}	128.89	Yes
G_3^{dh} vs G_4^{dh}	802.13	Yes
Between different pairs of comparisons for cursive		
G_3^{dc} vs G_4^{dc}	1.30e+03	Yes
Between hand-print and cursive of the same pair of comparisons		
$G_{3,4}^{sh}$ vs $G_{3,4}^{sc}$	1.38e+03	Yes
G_3^{dh} vs G_4^{dc}	2.08e+03	Yes
G_4^{dh} vs G_4^{dc}	565.86	Yes

FIGURES

Fig. 1

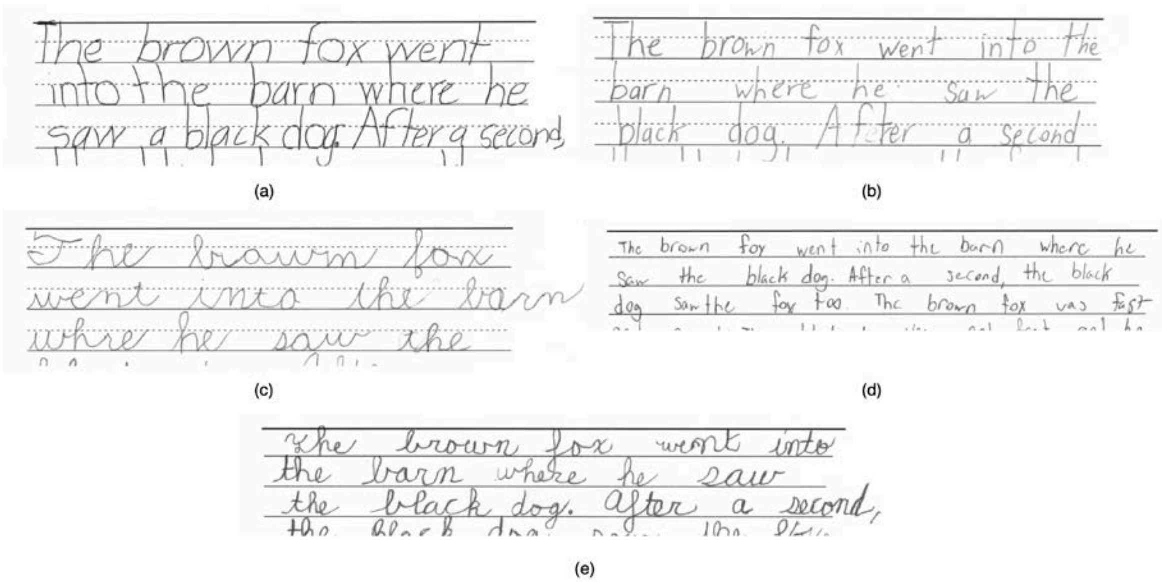


Fig. 2

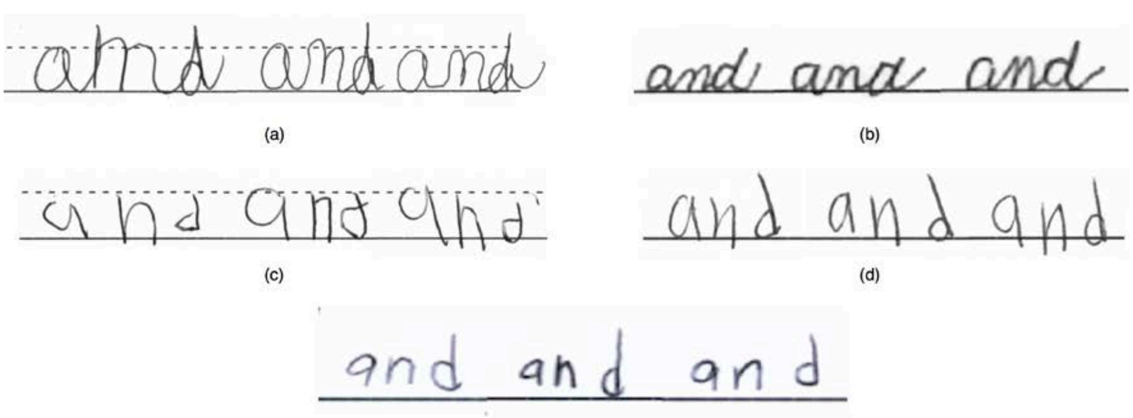
Fig. 3a

# strokes in "a"	one	two	three	uppercase	
formation of "a" staff	tented	retraced	looped	no staff	single down
# strokes in "n"	one	two	three	uppercase	
formation of "n" staff	tented	retraced	looped	no staff	single down
shape of arch of "n"	pointed	rounded			
# strokes in "d"	one	two	three	uppercase	
formation of "d" staff	tented	retraced	looped	no staff	single down
initial stroke of "d"	staff top	bulb			
Unusual formations	formation		symbol		
a-n relationship	a taller	a equal	a smaller		
a-d relationship	a taller	a equal	a smaller		
n-d relationship	n taller	n equal	n smaller		

Fig. 3b

Initial stroke of "a"	staff right	staff left	staff center	
Formation of "a" staff	tented	retraced	looped	no staff
Number of "n" arches	one	two		
Shape of "n" arches	pointed	rounded	retraced	combination
Location of "n" mid	above base	below base	at base	
Formation of "d" staff	tented	retraced	looped	
Formation of "d" initial	overhand	underhand	straight across	
Formation of "d" terminal	curved up	straight	curved down	no obvious end stroke
Symbol	unusual		symbol	
a-n relationship	a taller	a equal	a smaller	
a-d relationship	a taller	a equal	a smaller	
n-d relationship	n taller	n equal	n smaller	

Fig. 4



17

Fig. 5

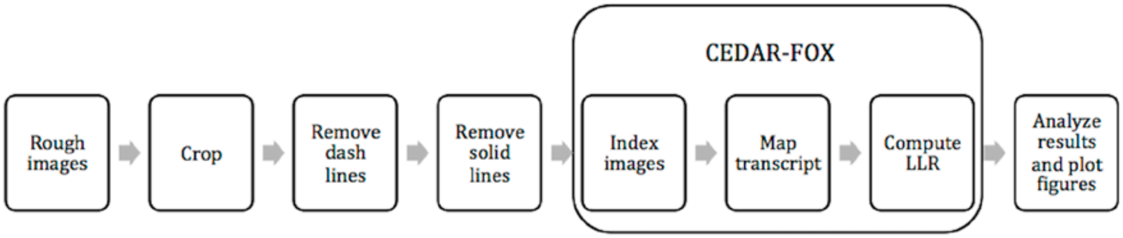

CEDAR-FOX
Rough images
Crop
Remove dash lines
Remove solid lines
Index images
Map transcript
Compute LLR
Analyze results and plot figures

Fig. 6

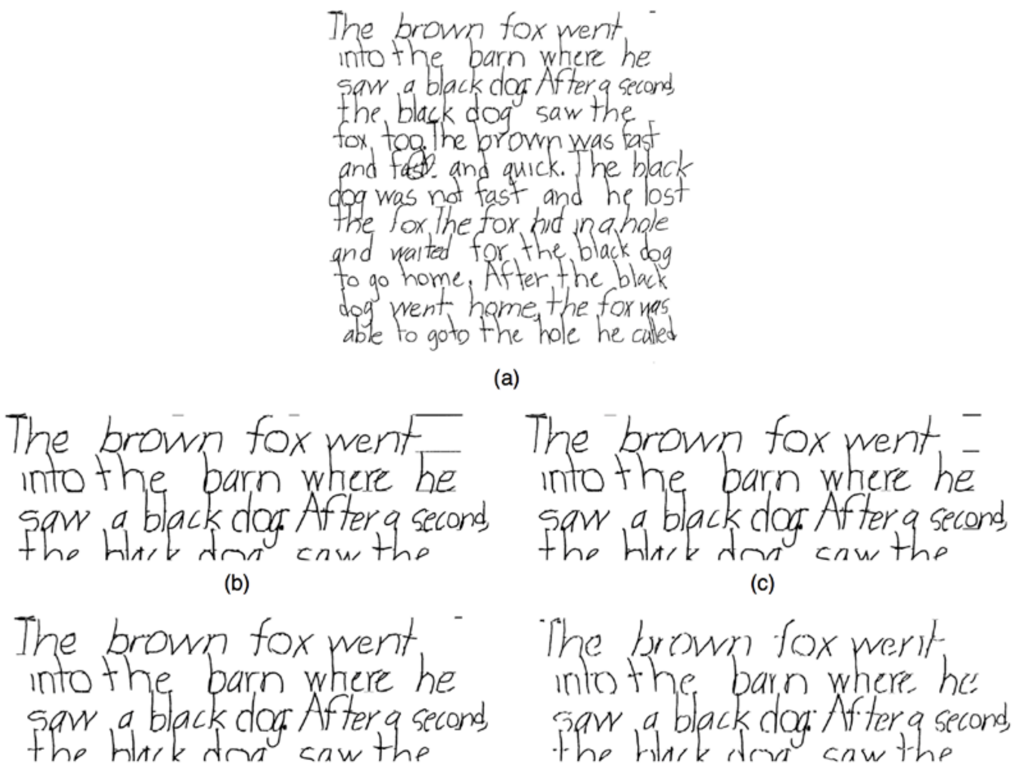

(a)
(b)
(c)
(d)
(e)

Fig. 7

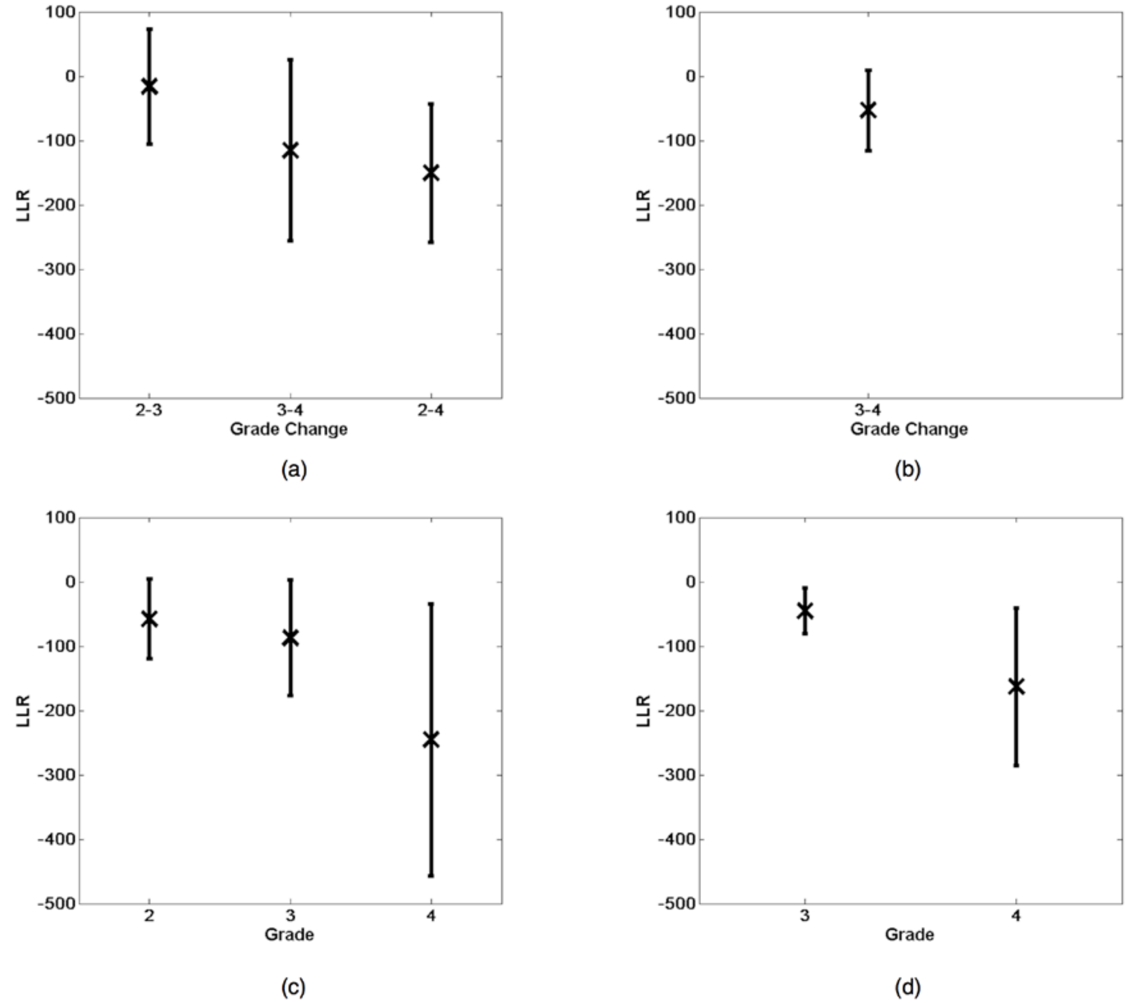

100
0
-100
-200
-300
-400
-500
LLR
2-3
3-4
2-4
Grade Change
(a)
3-4
Grade Change
(b)
2
3
4
Grade
(c)
3
4
Grade
(d)

Fig.83

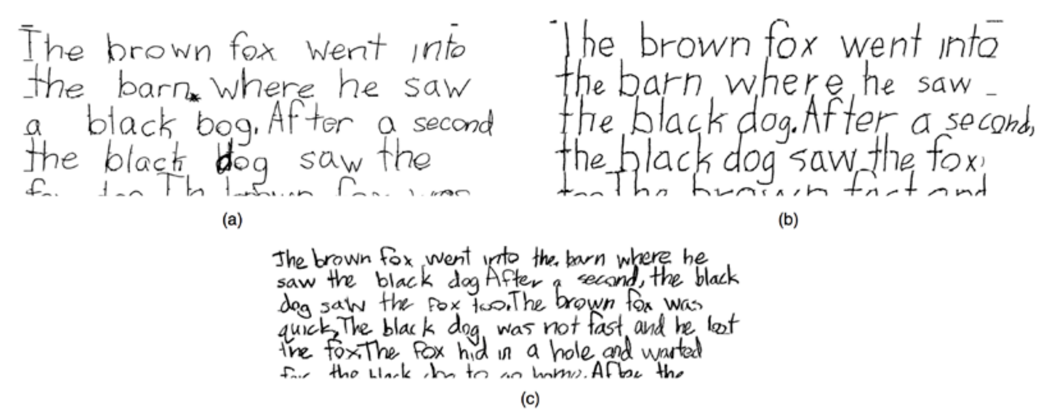
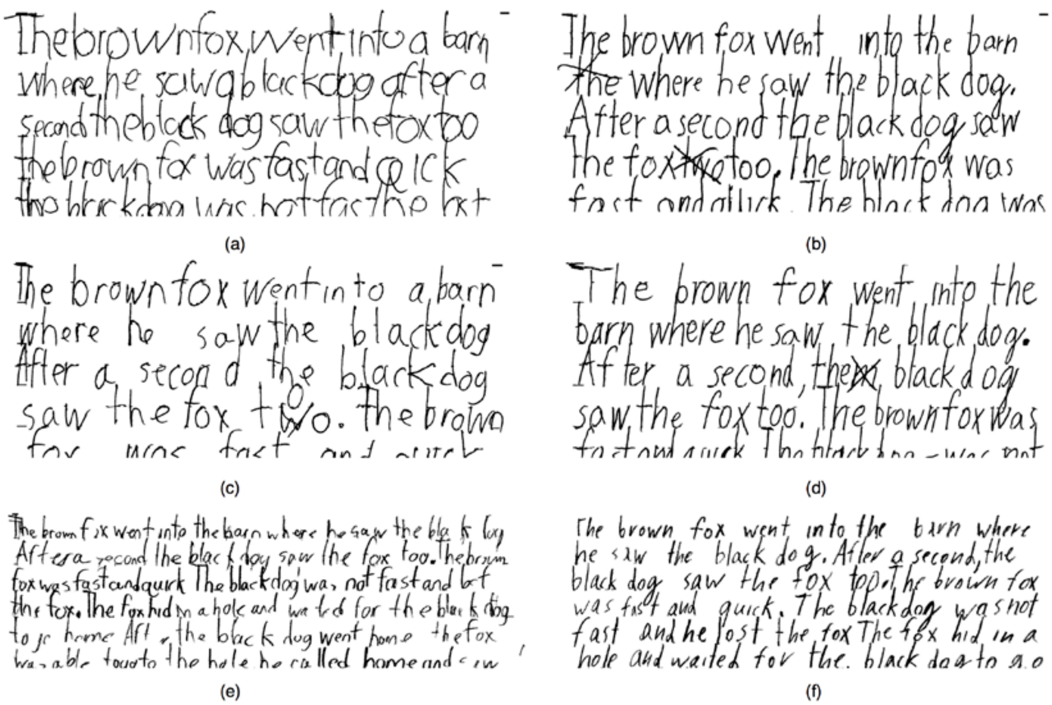

The brown fox went into
the barn where he saw
a black bog. After a second
the black dog saw the
(a)
The brown fox went into
the barn where he saw
the black dog. After a second,
the black dog saw the fox,
(b)
The brown fox went into the barn where he
saw the black dog After a second, the black
dog saw the fox too. The brown fox was
quick. The black dog was not fast and he lost
the fox. The fox hid in a hole and wanted
(c)

Fig. 9



Legends for illustrations

Fig. 1. Handwriting samples for a child: (a) hand printed in grade 2, (b) hand printed in grade 3, (c) cursive in grade 3, (d) hand printed in grade 4, and (e) cursive in grade 4.

Fig. 2. Zaner-Bloser copy book style for word *and*: (a) hand printed, and (b) cursive.

Fig. 3. Twelve characteristics of *and* together with their possible values: (a) hand printed and (b) cursive.

Fig. 4. Samples of Handwritten *and* together with feature values assigned by FDEs: (a) cursive grade 3, student ID: AlK2021WA11-12, features: 1, 2, 1, 1, 2, 1, 1, 0, 2, 0, 0, 0, (b) cursive grade 4, student ID: EmR2020Ll11-12, features: 2, 1, 1, 1, 2, 1, 0, 0, 2, 2, 2, 0, (c) hand printed grade 2, student ID: AlA2022RO11-12, features: 0, 0, 0, 1, 1, 0, 3, 0, 2, 0, 0, 0, (d) hand printed grade 3, student ID: EmL2021Za11-12, features: 0, 0, 0, 1, 0, 0, 1, 0, 2, 0, 0, 0, and (e) hand printed grade 4, student ID: AnL2020Fe11-12, features: 0, 1, 0, 1, 1, 0, 1, 0, 2, 0, 0, 0 .

Fig. 5. Automated comparison overview.

Fig. 6. Line removal with different parameter settings: (a) processed image, (b) #batch= 10, (c) #batch= 20, (d) #batch= 40, and (e) #batch= 80.

Fig. 7. Distributions of LLRs: (a) same child from different grades in hand printed, (b) same child from different grades in cursive, (c) different children in grades 2-4 in hand printed, and (d) different children in grades 3-4 in cursive.

Fig. 8. Hand printed samples for the same child in three grades: (a) grade 2 in 2011-12, (b) grade 3 in 2012-13, and (c) grade 4 in 2013-14. Between grades 2-3, LLR=51.71 indicating high similarity, while between grade 3-4 and 2-4, LLR=-105.15 and -148.05 indicating less similarity.

Fig. 9. Hand printed samples of two different children: (a, b) in grade 2 with LLR=-18.64, (c, d) in grade 3 with LLR=-101.99, and (e, f) in grade 4 with LLR=-178.74.

Scholarly Products Produced or In Process

1. S. N. Srihari, Z. Xu and L. Hanson, "Development of Handwriting Individuality: An Information-Theoretic Study", in *Proc. Int. Conf. Handwriting Recognition*, Crete, Greece, IEEE Computer Society Press, 2014.

2. S. N. Srihari, G. Chen, Z. Xu and L. Hanson, "Studies in Individuality: Can Students, Teachers and Schools be determined from Children's Handwriting?", *Proc. Int. Workshop on Computational Forensics*, Stockholm, Springer, Sweden, 2014.

3. M. Puri, S. N. Srihari and L. Hanson, "Probabilistic Modeling of Children's Handwriting", *CEDAR, University at Buffalo, The State University of New Your, Buffalo, NY, USA; Bureau of Criminal Apprehension Laboratory 1430 Maryland Ave. E. St. Paul, MN USA*

4. S. N. Srihari, and L. Hanson, "Development of Handwriting Individuality Among Children", *CEDAR, University at Buffalo, The State University of New Your, Buffalo, NY, USA; Bureau of Criminal Apprehension Laboratory 1430 Maryland Ave. E. St. Paul, MN USA*

5. Zhen Xu, Sargur N. Srihari and Lisa Hanson, "Minnesota Children's Handwriting Data with Missing Value Imputation", *CEDAR, University at Buffalo, The State University of New Your, Buffalo, NY, USA; Bureau of Criminal Apprehension Laboratory 1430 Maryland Ave. E. St. Paul, MN USA*

6. Zhen Xu, Sargur N. Srihari and Lisa Hanson, 'Minnesota Children's Handwriting Data: Features of word "and"', *CEDAR, University at Buffalo, The State University of New Your, Buffalo, NY, USA; Bureau of Criminal Apprehension Laboratory 1430 Maryland Ave. E. St. Paul, MN USA*