A Tablet-Based App to Discriminate Children at Potential Risk of Handwriting Alterations in a Preliteracy Stage

Linda G. Dui, Francesca Lunardini, Cristiano Termine, Matteo Matteucci, Simona Ferrante

Abstract— **Failing to master handwriting, as in the case of Dysgraphia, has negative consequences on children's lives. In early stage of development, Dysgraphia diagnosis is delayed and not easily achievable. Thus, the aim of this work is to propose a valid tool to anticipate Dysgraphia screening at a preliteracy age. We developed a tablet application to analyze characteristics altered in dysgraphic handwriting, such as rhythmical laws (isochrony and homothety), or a collection of kinematic and dynamic parameters (smoothness, pressure, frequency contents). To be suitable for the pre-literacy stage, possible alterations are investigated in symbol drawings. The app is tested on 104 preschoolers, both with normal (n=76) and delayed graphical abilities (n=28), reporting excellent acceptance. Some isochrony alterations were reported only for children with delayed graphical abilities. Moreover, kinematic and dynamic parameters are effective in discriminating between risk and no-risk conditions. Indeed, the logistic classification adopted resulted in a 0.819 area under the precision-recall curve. These findings pave the way toward an early screening of future handwriting alteration, starting from a pre-literacy age.**

Clinical Relevance — **The analysis of symbol drawing on a tablet-based app can be used to discriminate kindergartners at risk of handwriting alterations.**

I. Introduction

Dysgraphia affects 5% to 27% of school population [1], [2], with negative consequences on several aspects of children's lives [3]. Children with dysgraphia have problems in learning handwriting, a fundamental ability usually mastered by the third year of primary school [4]. Thus, it is almost impossible to diagnose dysgraphia before that age, and to distinguish it from a simple developmental delay, which would likely recover after the empowerment of weak abilities through dedicated didactical activities, or a specific occupational therapy training. When a diagnosis is finally possible, neuropsychiatrists' offices are overloaded also with such non-severe cases, with increasing social costs. Moreover, children living in disadvantaged social conditions might never reach specialists' office for a diagnosis.

To support in the screening for handwriting abnormalities, schools are implementing observational programs starting form the last year of kindergarten. Such programs are aimed at studying skills correlated to handwriting fundamentals, such as spatial awareness, eye-hand coordination, and dexterity [3], with the final goal of intervening with specific exercises in case of abnormalities. Even though experienced teachers can be a first aid, simple observation cannot disclose small yet relevant differences among children, nor accurately track their evolution.

Therefore, there is the strong need for innovative solutions for the early screening of handwriting weaknesses. Such tools should meet specific requirements: first of all, the ability to investigate characteristics typically altered in dysgraphic handwriting and distinguish them from a mere poorly handwritten text. Second, to anticipate the diagnosis, they must investigate such characteristics in a pre-literacy stage. Third, to broaden the screening to preclinical environments, such tools should be easily accessible to nonclinical users, such as teachers. Fourth, gamification or engaging features should be leveraged to increase children's engagement and boost system acceptance.

For a quantitative assessment of handwriting, first, we considered two features found to be altered when dysgraphia is present: isochrony and homothety [5]. Isochrony predicts that writing speed is increased when size is increased, to keep execution time approximately constant [6]. Homothety predicts that the fraction of time dedicated to each letter, with respect to the total word time, is constant, independently from writing size [6]. These laws, together, assure constancy in absolute and relative time to write letters in words, at different writing sizes. Second, we considered a collection of parameters proper of handwriting kinematics and dynamics, such as fluidity, frequency content, or pressure, which together are proven to discriminate between normal and dysgraphic handwrites [7], [8].

To anticipate the identification of writing abnormalities to preliteracy stages, we study such laws on symbols.

To widen the screening to non-clinical environments (e.g., home, school), we propose the use of commercial tablets, which additionally allow for gamification, to further involve children in longitudinal monitoring.

This study has two goals: (i) to test children's acceptance of a new tablet application, designed to satisfy the aforementioned needs; (ii) to identify the effectiveness of objective parameters in detecting children considered at risk based on teachers' observation.

II. Methods

A. Material

We developed *Play-Draw-Write*, an application in Unity 2018.3.2f1, for an iPad 6, with Apple Pencil 1 (pen position and pressure on the screen sampled at 240 Hz). The app is designed to study the isochrony and homothety laws on symbols. The app interface presents an empty canvas, with the example of the symbol (or sequence of symbols) to copy, i.e., a square, or a sequence of a circle, a line, and a reversed U (Fig. 1). The app also suggests the writing modality (spontaneous, big, or small), so that it is possible to test for both speed changes due to size modulation, and for fraction time differences between modalities for each symbol of the sequence. A section of the app is dedicated to a satisfaction questionnaire. The 5-point likert scale questions are:

1. Was the pen comfortable to use?
2. Was the pen light?
3. Do you prefer this pen rather than those you always use?
4. Did you enjoy the game and do you wish to continue using it at home?
5. Are you satisfied with the experience?

The likert scale was paired with emoticons, to help children indicate the preferred answer.

B. Participants and Protocol

Children from the last year of kindergarten, without known pathologies, were enrolled in the study. Trained teachers expressed their opinion on children's graphical abilities, according to a checklist provided by neuropsychiatrists. Children were categorized into a Risk group (R) and a Typically Developing group (TD) accordingly.

The protocol included three activities, executed once per child, as follows:

1. Copy the square in three modalities: spontaneous, big, and small;
2. Copy the sequence of symbols in three modalities: spontaneous, big, and small;
3. Answer the questionnaire.

The protocol was approved by the Politecnico di Milano Ethical Committee n. 24/2019, and informed consent was signed by participants' parents.

C. Statistical analysis

Non-parametric tests were used, after checking for normality with a Lilliefors test.

To confirm that children were compliant with the instruction of changing drawing size, we checked for significant differences in tract length between the three modalities, by means of a Friedman test and Wilcoxon matched pairs post hoc with Bonferroni correction.

To test for isochrony, we considered both activities 1 and 2. We computed the instantaneous speed as the derivative of the distance between two successive points, smoothed with a 5 Hz lowpass filter, and we averaged it on the entire execution. We tested for significant differences between modalities with a Friedman test and post hoc in both groups, as we expected a modulation of the average speed proportional to size. Finally, to check for between-group difference, we performed a Mann-Whitney test between length and speed for each modality.

To test for homothety, we considered the time devoted to each symbol in the sequence with respect to the total sequence (fraction time) in activity 2. We looked for differences in fraction time between writing modalities through a Friedman test. If significance was reached, Bonferroni-corrected post hoc was executed.

Figure 1. Interface to copy the sequence of symbols in the spontaneous modality.

To further analyze if drawing kinematics and dynamics had the ability to discriminate between the two groups, we combined parameters related to:

- specific exercise requirements (tract length, between-modality difference in tract length, and speed);
- gesture smoothness (signal to noise velocity peaks [9], number of pen-up events, number of stops (speed < 0.5cm/s) on the screen, stop time/total time ratio [10]);
- pressure (mean value, dominant frequency);
- drawing strategy (starting point and direction for drawing the square).

We enriched the feature set with children characteristics, such as, age, gender, handedness, and technology acquaintance. To check that the risk class was not biased by age, we performed a Mann-Whitney test between the R and TD groups.

We binarized the drawing direction and handedness with the one-hot encoding technique [11], whilst numerical features were standardized. We split the dataset into a training and a holdout test set. We selected the most important features from the training set with a forward stepwise correlation-based feature subset selection [12]. We applied a ten-fold cross-validation, stratified on risk basis, and we retained features selected at least in one fold.

Then, among machine learning models, we chose a logistic classifier to test the predictive power of these features in discriminating the risk condition. We trained it with a ten-fold cross-validation stratified on risk, and considered the test set weighted average of the area under the precision-recall curve (AUPRC, x-axis: recall, y-axis: precision), which is suitable in cases of heavy class imbalance [13].

As for the questionnaire, we performed frequency analysis of the five questions.

Data processing was executed in Matlab R2018b, statistical analysis was performed in R 3.3.3 (significance 5%), feature selection and classification were performed in Weka 3.8.

III. Results

104 children participated in the study. Table 1 summarizes population statistics, for each group. We confirmed that the risk condition was not caused by age difference (Mann-Whitney test: p=0.413).

TABLE I. POPULATION STATISTICS. R = RIGHT, L = LEFT, Y = YEARS, M = MONTHS, ACQ. = ACQUAINTANCE, NA = NOT AVAILABLE

	At Risk	Typically Dev.
Gender	22 males, 6 females	37 males, 39 fem.
Handedness	23 R, 2 L, 3 both	65 R, 8 L, 3 both
Age	5 y 6 m ± 2 m	5 y 6 m ± 3 m
Tablet acq.	14 yes, 13 no, 1 NA	46 yes, 29 no
Stylus acq.	3 yes, 24 no, 1 NA	6 yes, 69 no

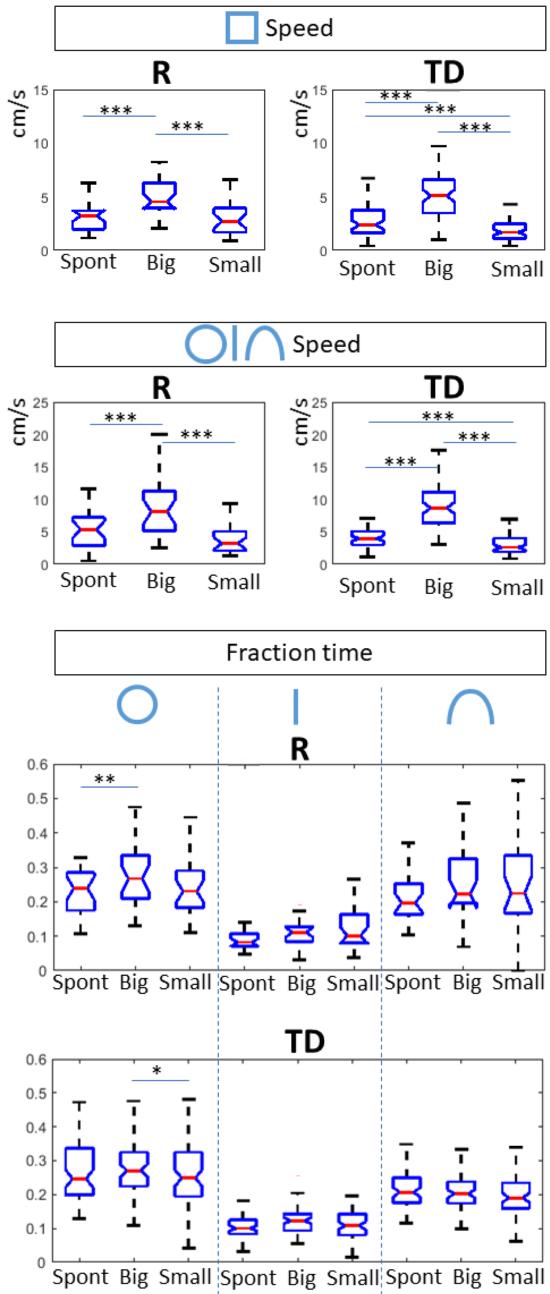

Figure 2. Isochrony and homothety results, divided by R and TD children. Isochrony: speed (cm/s) comparison between modalities. Homothety: fraction time (normalized units) comparison between modalities. Horizontal red lines are medians, boxes are inter-quartile ranges, notches are 95% confidence interval for the median. Asterisks: * $p<0.05$; ** $p<0.005$; *** $p<0.001$.

As for activity 1 and 2, size modulation resulted in significant differences in both groups ($p<0.001$ for both the square and the sequence). Post hoc revealed that significance was always reached ($p\leq 0.002$). The comparison between the two groups in terms of tract length revealed that R children spontaneous copy cannot be distinguished by the TD group when drawing the square (p=0.217), but it is significantly longer when drawing the sequence (p=0.003). The big modality length was not significantly different in the two groups (square: p=0.977, sequence: p=0.528), but the small modality was again longer for R children (square: p=0.003, sequence: p=0.003).

As for isochrony (Fig. 2 – Speed), Friedman test revealed that significant differences in average speed between writing modalities exist in both groups, for both exercises ($p<0.001$). In the post hoc, the TD group showed significant differences between each pair of modalities ($p<0.001$). On the counterpart, the R group did not reach significance in the post hoc for the spontaneous-small couple (square: p=1, sequence: p=0.050).

Between-group difference in speed emerged for the small square only (p=0.002), with a faster execution observed in the R group. As for homothety (Fig. 2 – Fraction time), the TD group respected homothety in the majority of cases, as the fraction time devoted to each symbol did not differ according to the writing modality in lines (p=0.100) and reversed U (p=0.087), but Friedman test resulted in p=0.050 for the circle, and the post hoc confirmed a significant difference in the big-small couple (p=0.020), but not in the other pairs (spont.-big: p=0.496, spont.-small: p=0.143). The R group did not violate homothety in lines (0.077) and reversed U (p=0.527), but the Friedman test reported a significant effect of modality on the fraction time for circles (p=0.006). The post hoc revealed a significant difference in the spontaneous-big couple (p=0.004), whilst the other couples did not show significant differences (spont.-small: p=1, big-small: p=0.407)

As for the risk discrimination, 9 children were excluded, as they lacked a complete predictors' set. The hold out test set comprised 8 R and 37 TD children. The most important features for the prediction are reported in Table 2, with importance expressed as the percentage of folds they were selected in. Important features did not comprise technology acquaintance, thus suggesting that it did not biased the result. The logistic classification resulted in a 0.819 AUPRC.

The questionnaire revealed a prevalence of positive answers to all the items. Fig. 3 reports the percentage of the answer to each question.

TABLE II. IMPORTANT FEATURES

Feature	**Importance**
Drawing strategy (square)	90%
Gender	40%
Dominant frequency (big square)	20%
Size (spontaneous square)	10%
Size variation (sequence)	10%
Mean speed (small square, spont. sequence)	10%
Mean acceleration (small square)	10%

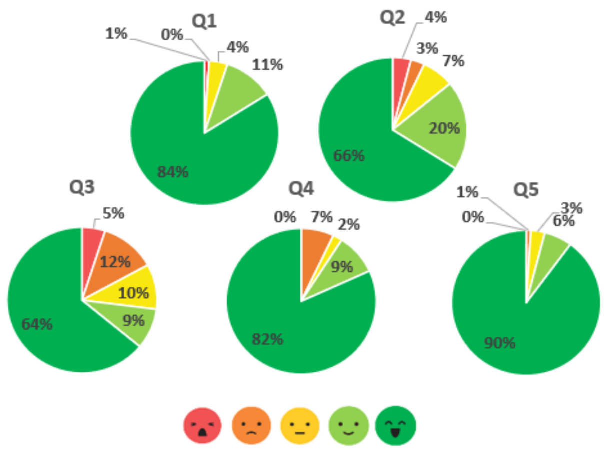

Figure 3. Questionnaire answers. Q1 to Q5 represent the five questions. Answers are reported in pie charts, colored according to the emoticons in the interface (dark green: most positive, red: most negative).

IV. Discussion

In this study, we present *Play-Draw-Write*, a tablet-based application designed to anticipate the screening for handwriting problems to an age when handwriting is not learned yet. As dysgraphic handwriting shows alteration in several aspect of gesture production, such as rhythmical components (isochrony and homothety [5]), but also pressure, frequency content or smoothness [7], [8], we here show that the same alterations can potentially be detected in preliteracy symbol drawings. To this end, we tested a group of preschoolers considered at risk of delay in graphical abilities (R), according to their teachers' judgement, and we compared their performance to typically developing (TD) peers.

R children showed a violation of the isochrony principle, as they do not always adapt their speed to changes in size. This cannot be ascribed to a difficulty to comply with the instruction of varying the drawing size itself, as changes in tract length were always significant. Such results show that it is possible to anticipate the evaluation of isochrony alterations, typical of Dysgraphia, to an age when handwriting is not learned yet.

As for homothety, we found a single exception in both groups, which cannot completely undermine its validity in none of the groups. However, the relatively small number of R children is reflected in wider confidence interval for fraction time medians, thus making the violation in the R group more significant.

Concerning gesture production kinematics, results from the classification are promising. Indeed, contrary to previous studies where drawings did not seem to be effective in groups discrimination [5], our findings support the hypothesis that handwriting and symbol drawing might share common characteristics and difficulties. To further refine the classification and reach higher performance, we plan to broaden our sample size and to include new features in the model.

A limitation of the study is that the stratification between R and TD children is based on the important, yet subjective and nonclinical, teachers' judgment, and it is not possible to confirm that real handwriting difficulties would arise in the following years. To assure that inconsistencies in results are not due to possible misclassification, a long-term follow-up is needed.

Questionnaire results reflected the enthusiastic reaction observed in children during game execution. They felt comfortable with the new writing tool and they liked the overall experience. This is encouraging, in a perspective of repeating the test over time to monitor children's evolution, without boring them.

To conclude, we provide a new tool able to detect alterations - typical of Dysgraphia handwriting - also in symbol drawing. This is a fundamental step towards the early diagnosis of this Learning Disability, which will enable more targeted interventions and improve the whole children's lives.

Acknowledgments

We would like to thank "Ministero dell'Istruzione, Ufficio Scolastico Regionale per la Lombardia Ufficio XIV – Varese" for granting schools availability.

References

[1] R. Karlsdottir and T. Stefansson, "Problems in Developing Functional Handwriting," *Percept. Mot. Skills*, vol. 94, no. 2, pp. 623–662, 2002.

[2] M. Al-Yagon *et al.*, " The Proposed Changes for DSM-5 for SLD and ADHD ," *J. Learn. Disabil.*, vol. 46, no. 1, pp. 58–72, 2012.

[3] K. P. Feder and A. Majnemer, "Handwriting development, competency, and intervention," *Developmental Medicine and Child Neurology*, vol. 49, no. 4. pp. 312–317, 2007.

[4] A. Overvelde and W. Hulstijn, "Handwriting development in grade 2 and grade 3 primary school children with normal, at risk, or dysgraphic characteristics," *Res. Dev. Disabil.*, vol. 32, no. 2, pp. 540–548, 2011.

[5] E. Pagliarini *et al.*, "Dyslexic children fail to comply with the rhythmic constraints of handwriting," *Hum. Mov. Sci.*, vol. 42, pp. 161–182, 2015.

[6] P. Viviani and C. Terzuolo, "Trajectory determines movement dynamics," *Neuroscience*, vol. 7, no. 2, pp. 431–437, 1982.

[7] T. Asselborn *et al.*, "Automated human-level diagnosis of dysgraphia using a consumer tablet," *npj Digit. Med.*, vol. 1, no. 1, 2018.

[8] J. Mekyska, M. Faundez-Zanuy, Z. Mzourek, Z. Galaz, Z. Smekal, and S. Rosenblum, "Identification and Rating of Developmental Dysgraphia by Handwriting Analysis," *IEEE Trans. Human-Machine Syst.*, vol. 47, no. 2, pp. 235–248, 2017.

[9] J. Danna, V. Paz-Villagrán, and J. L. Velay, "Signal-to-Noise velocity peaks difference: A new method for evaluating the handwriting movement fluency in children with dysgraphia," *Res. Dev. Disabil.*, 2013.

[10] M. M. Prunty, A. L. Barnett, K. Wilmut, and M. S. Plumb, "Handwriting speed in children with Developmental Coordination Disorder: Are they really slower?," *Res. Dev. Disabil.*, vol. 34, no. 9, pp. 2927–2936, 2013.

[11] A. Ghatak, *Machine Learning with R*. 2017.

[12] M. a. Hall and L. a. Smith, "Practical feature subset selection for machine learning," *Comput. Sci.*, 1998.

[13] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in *ACM International Conference Proceeding Series*, 2006.