

# Classification of Children's Handwriting Errors for the Design of an Educational Co-Writer Robotic Peer

Shruti Chandra<sup>†,\*</sup>, Pierre Dillenbourg<sup>†</sup> and Ana Paiva<sup>\*</sup>

<sup>†</sup>CHILI, École Polytechnique Fédérale de Lausanne, Switzerland

<sup>\*</sup>INESC-ID & Instituto Superior Técnico, Universidade de Lisboa, Portugal

shruti.chandra@tecnico.ulisboa.pt, pierre.dillenbourg@epfl.ch, ana.paiva@inesc-id.pt

## ABSTRACT

In this paper, we propose a taxonomy of handwriting errors exhibited by children as a way to build adequate strategies for integration with a co-writing peer. The exploration includes the collection of letters written by children in an initial study, which were then revised in a second study. The second study also analyses the “peer-learning” (PL) and “peer-tutoring” (PT) learning methods in an educational scenario, where a pair of children perform a collaborative writing activity in the presence of a robot facilitator. The data obtained in the first two studies allowed us to create a “taxonomy of handwriting errors”. A set of writing errors were selected and implemented in an educational activity for validation. This activity constituted a third study, wherein we systematically induced the errors into a Nao robot's handwriting using the PT method - A teacher-child corrects the handwriting errors of the learner-robot. The preliminary results suggest that the children in general showed awareness to the writing errors and were able to perceive the writing abilities of the robot.

## ACM Classification Keywords

K.3.1. Computer Uses in Education: Collaborative learning;  
K.3.2. Computer and Information Science Education: Information systems education

## Author Keywords

Peer assisted learning methods; handwriting difficulties; child-robot-interaction;

## INTRODUCTION

Handwriting problems in children affect them negatively with respect to both their academic performance [13] and their motivation and self-esteem [28]. The development of young children's writing ability is not only considered an essential criterion for success in school [32, 35] but also a critical skill that will impact all of us throughout adulthood [15]. A few studies have indicated that handwriting can often be used to judge children in schools; students with poor handwriting

are assigned lower grades compared to those with legible handwriting despite similar content [12, 36]. The number of children presenting handwriting problems is estimated to be between 10 and 30% [24], and their consequences include an inability to keep up with the written work required in the class, to obtain high scores in Mathematics, and to achieve sustained attention [31, 26]. Quite often, children with handwriting problems are mislabeled as lazy, which in turn causes additional frustration and behavioral problems [31].

Mastering handwriting involves a complex blend of motor-perceptual abilities, and ineffective motor skills are difficult to change once they are acquired. Hence, special attention should be given to sharpening the motor and cognition skills of preschoolers from the very beginning [30] to prevent problems later on.

Research in educational robotics has proven its significance in different fields such as mathematics, programming, science, games and design[20]. During the past few years, social robots have been used with children to explore new methods of learning in education. For example, Kanda et al. used a humanoid robot as an English peer-tutor for Japanese students [22] in a study. The results showed that the robot encouraged children to improve their English vocabulary and was able to form relationships with them. Furthermore, the EMOTE project<sup>1</sup> addressed the role of empathy in robots [10].

The current work is conducted under the *CoWriter project*<sup>2</sup>, which explores the long-term research question of *how to create a social robot that can help children to acquire handwriting skills*. With this vision in mind, we have been investigating the following aspects:

**Handling writing difficulties** - to explore children's common handwriting errors, which may differ according to their age and motor-perceptual abilities.

**Modes of Interaction** - to determine the more adequate modes of child-robot interaction to maximise the learning gains of children having different levels of writing skills.

“Peer-learning”(PL) and “peer-tutoring”(PT) are two types of peer-assisted learning methods effectively used in the field of education. In both these methods, the interaction generally occurs between two students. In the PL method,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

IDC '17 June 27-30, 2017, Stanford, CA, USA

© 2017 ACM. ISBN 978-1-4503-4921-5/17/06...\$15.00

DOI: <http://dx.doi.org/10.1145/3078072.3079750>

<sup>1</sup>[www.emote-project.eu](http://www.emote-project.eu)

<sup>2</sup><http://chili.epfl.ch/cowriter>

there is an equal alliance between the two students, whereas in the PT method, one student holds responsibility over another student [25].

**Learning benefits through child-robot interaction** - to use a robot as an agent for enhancing children's learning. In this interaction, it becomes necessary to understand how children perceive the abilities and behavior of a social robot that was endowed with not only writing competence but also writing difficulties.

This article investigates the above-mentioned aspects through a series of three studies. In these studies, we show how a taxonomy of writing difficulties in letters, more specifically, the English alphabet, can be exploited to effectively build a robotic peer to help children. The taxonomy can be used by an automated system, such as a social robot; school teachers; or parents to understand a particular writing issue faced by a child and then help him/her to correct that particular issue.

Section "*Study I*" explains a first study that was conducted with children 4-6 years of age to collect their handwritten letters to identify common handwriting errors and to test learning modes of interactions. Section "*Study II*" presents a second study with children 6-8 years of age to investigate further writing errors to revise the built taxonomy, extend it, and explore the modes of interaction in the presence of a social robot. Finally, section "*Study III*" details the implementation of a set of three writing errors from the taxonomy in a system for the validation and testing of the errors in a third study. The study combined the results of the first and second studies into a version of a robotic peer that has difficulties and needs to learn from the child. When introducing a robotic peer, it is crucial to understand how children perceive the robot's abilities and behavior in a child-robot interaction. Therefore, in this final study, we also explored the children's perception of a "Learning" vs "Non-learning" robot. Finally, the last section concludes by reiterating the contributions of the article.

## BACKGROUND

It is believed that the development of handwriting in children begins with early scribbling on paper or other surfaces after they are able to grasp tools for handwriting or drawing. The literature also suggests that writing instructions should be given until a child masters nine shapes, including lines, circle, triangles and oblique crosses [6]. At the ages of 2-3, a child begins to form basic geometric shapes, such as horizontal and vertical lines, followed by being able to imitate a cross, square and triangle between the ages of 4-6 years [6]. The ability to copy the shape of an oblique cross is seen as readiness for handwriting [6, 2]. The research indicates that the typical time period to develop the quality of handwriting lies between primary school grades 1-5 (ages 6-11). Among these years, the level of skill improves more quickly during Grade 1 (ages 6-7) and reaches a plateau by Grade 2 (ages 7-8). Further development during the ages of 8-9 leads to the emergence of more automatic and organised letter formation, which is regarded as a foundation for fluent writing [24, 8].

Handwriting is a complex motor-perceptual ability and includes several components: coordination of visual-motor

skills, perceptual and cognitive skills, motor planning, and kinesthetic and tactile sensitivities [27]. In addition, external factors related to environment/biochemical issues, such as sitting position, posture, and chair/desk height, may also influence the handwriting performance [15]. Because handwriting is a combination of forms of letter (orthographic codes), names of letter (phonological codes) and shapes of letters (graphomotor codes) [5], researchers have argued that memory, more specifically the ability to recall the shapes of letters and orthographic processes, is more responsible for handwriting skills rather than motor skills [4].

Simner [33] discussed that the writing errors produced by kindergarten children include additions, deletions and misalignment in parts of a letter. These errors not only influence their academic performance in the kindergarten period but also throughout their first grade. Other common writing errors, due to insufficient motor skills often observed in grade 1 children (ages 6-7 years), are incorrect sizes of letters, placement of letters, and relationships among parts in a letter [34]. The motor skill aspects include isolation, grading and timing of movements. The isolation and grading aspects resulted from the inadequate grasping of the pen. In addition, slow and haphazard writing are associated with the timing of movements affecting the flow of writing [14]. The result of a survey with 169 primary grade teachers indicated that 23% (SD = 14%) of the children in their class faced difficulty with handwriting and that the most common handwriting problems experienced by students are related to overall neatness (76%), spacing between words (66%), letter size (59%), letter formation (57%), alignment of letters (54%), and reversals (52%) [16].

In this article, we focused on classifying the common handwriting errors of children 4-8 years of age because of the lack of extensive details on the writing errors.

In the field of education, researchers have used peer-assisted learning methods either involving an interaction between students or between a robot and a student to support childhood learning. The PL has a reciprocal learning style, where both students act as both teacher and learner [9] without holding power over each other [9, 25]. In a two-month field study, Kanda et al. used Robovie [23], a humanoid robot as a peer-learner/friend, and the results indicated that the robot was able to make friendly relations with children. On the other hand, the PT learning method (also known as the learning-by-teaching method) relies on a one-way learning style: there is an unambiguous distinction between the learning and teaching roles among the students [9]. One student acts as a teacher, while the other one acts as a learner. The research also indicated that when a person teaches someone, knowledge becomes clearer and better organised in his/her mind because of the feedback and communication of ideas [7]. Tanaka et al [37] used the PT method in a study wherein children taught English verbs to a Nao robot<sup>3</sup>, which acted as a care receiving interactive agent. The results suggested that the robot was able to enhance the children's spontaneous learning and accelerated their motivation to provide care-taking.

---

<sup>3</sup>Aldebaran robotics: <https://www.aldebaran.com/en>.

To design better robot-aided learning scenarios for children, it is crucial to understand how children perceive the abilities and behavior of a robot. In a study by T.N. Beran et al., conducted with 168 students, the results have shown that a significant number of students attributed to the robot cognitive and behavioral characteristics [3]. Likewise, another study by Kahn et al., indicated that the majority of children ascribed a robot with mental states and as a social being [21]. In a similar line of research, we explored children’s perception of the abilities and behavior of a *Learning vs. Non-Learning* robot.

We used the PL and PT learning methods as modes of interaction to explore their effectiveness in the acquisition of handwriting skills. Recent studies have also explored these methods involving a child-robot interaction to improve children’s ability in forming shapes and handwriting. For example, Shizuko et al. [29] performed a study with a tele-operated Nao robot whereby children drew some shapes associated with English words. The results suggested that the robot helped children to learn unknown English words. Following a similar line, Hood et al. [18] used the PT method in a learning scenario whereby children taught handwriting skills to the autonomous Nao robot. The results indicated greater engagement with the robot by the children and their improved writing skills. Another long-session study was conducted with children with handwriting difficulty using the PT method, where the Nao robot adapted to the children’s handwriting issues. The results indicated the children’s increased self-confidence and commitment towards the robot [19].

Because the previous studies have introduced random writing errors, we attempt to provide a classification of errors that will allow them to be approached in a systematic manner. The work presented in the article makes two major contributions: 1) Proposal of a taxonomy of writing issues present in children’s handwritten letters in the age group of 4-8 years - which may be suitable for the educational community in evaluating children’s handwriting skills. 2) Test of a set of writing errors from the built taxonomy in an educational scenario in a 4-week-long study. The goal of the last study is to understand children’s perception towards a robot having writing difficulties and showing two different competencies: ‘*Learning*’ vs. ‘*Non-Learning*’.

## DESIGNING MODES OF INTERACTIONS

In the context of this work, we attempted to develop educational activities that are intended to help children in the acquisition of handwriting skills. Because the research targeted young children, it was crucial to design an activity that was playful, simple to understand, effective and capable of sustaining children’s engagement for a certain period of time. We chose the PL and PT learning methods as modes of interactions to understand different aspects of children’s behavior and writing skills. The two designed interaction modes were termed as “peer-learner mode” and “teacher-learner mode” and served as two conditions for the first study. Each interaction mode involved a pair of children with a human facilitator, and a collaborative writing activity was performed. The role



Figure 1: Children using: (a) Paper; (b) Tablet

of the facilitator was to provide the instructions to the children during the activity. The modes are explained below:

- **Peer-learner mode:** In this mode, no role was assigned to the children. The experimenter showed a letter card to both children and asked them to copy the letter onto their respective sheets. After they finished writing, the experimenter prompted them to provide corrections on each other’s writing performance. The activity was repeated with other letter cards. For half of the letters, children used a paper sheet, whereas for the other half, they used a tactile device (one per child).
- **Teacher-learner mode:** In this mode, the experimenter assigned roles to children: one child acted as a learner, and the other acted as a teacher. The experimenter asked the teacher-child to show a letter card to the learner-child so that s/he could copy the letter onto a paper sheet. Then, the experimenter prompted the teacher-child to provide corrective feedback on the learner-child’s writing performance. Here, the children again used both paper sheets and tactile devices. After copying half of the letters, the experimenter exchanged their roles, and the activity was repeated with the remaining letters.

## STUDY I

### Objectives

The overall goal of the study was to form a database of children’s handwritten letters to explore the most common handwriting errors. In addition, the other objectives were as follows:

- To identify the common writing errors in the letter samples of the collected data during the study and to build a taxonomy of these errors.
- To assess the effectiveness of both interaction modes.

### Methodology

We conducted a first study with 20 English-speaking children in the age group of 4 to 6 years (reception class) in a school. The study followed the ethical norms of privacy and responsibility. As such, only children who assented to the study and whose parents signed the consent form participated. The study consisted of a between-subject design, and out of 20 children, 10 children (5 pairs) participated in the peer-learner mode ( $M = 4.95$ ,  $SD = .15$ ; 6 male and 4 female), and the other 10 participated in the teacher-learner mode ( $M = 4.95$ ,  $SD = .41$ ; 3 male and 7 female). The materials used in the study were an experimental record sheet, paper sheets, tablets

(tactile devices), tablet pens, a camera with a microphone, and target cards with printed letters. For the pre- and post-test, we chose 8 letters, a combination of upper- and lowercase letters ( $\alpha$ ,  $\pi$ , y, j, S, N, m and Q) printed on 8 separate cards and based on different geometrical shapes. Out of the 8 letters, 6 letters ( $\alpha$ ,  $\pi$ , y, j, S and Q) were included in the main writing activity. The  $\alpha$  and  $\pi$  letters were only included among other English letters to observe the transfer skill, as we knew that children did not have prior knowledge of these Greek letters. For the tactile device, a specific writing app was developed to provide the dual-user interactive feature: when a shape is drawn on one tablet, it can be observed on the other tablet in real time. In addition, an eraser button is provided for children to erase their own writing. This functionality allowed a pair of children to correct each other's writing on their own tablet in real time. The interaction modes were presented as a game to children, and each interaction lasted approximately 10-14 minutes. The study followed the phases detailed below:

**Pre-test:** In this phase, an experimenter displayed the 8 target letter cards on the table to a pair of children and asked them to write the letters individually on a paper of sheet.

**Writing activity:** After finishing the pre-test, the experimenter performed the writing activity with 6 target letter cards according to one of the interaction modes selected for the children (see Fig. 1) (explained in section- *Designing modes of interaction*).

**Post-test:** After finishing the writing activity, the pre-test activity was repeated and also served as a post-test.

## Analysis

To analyse the handwriting errors in the letters, we collected the data from the paper sheets as well as from the tablets. The data were analysed by two independent coders, who examined each sample by comparing it with the letter target card (letter presented during the study). Then, the samples were classified based on various factors such as the shapes of the trajectory and sub-trajectory, the number of strokes, and the alignment. Furthermore, a taxonomy of writing issues was built depending on the classification. To capture the communication modalities, two other coders annotated the videos using the ELAN multimedia annotation tool<sup>4</sup>. In terms of the reliability of a participant's writing errors and communication modalities, Cohen's kappa showed agreement values of .65 and .85, indicating a good agreement.

## Results

To better understand the interactions among children, we classified the modalities into three main categories: verbal (*e.g.* explanations/short answers), gesture (*e.g.* Pointing), and back channel (*e.g.* facial expressions, nodding etc). Due to the small size of the sample data, the results are discussed in terms of frequency. In terms of corrective feedback, we observed that the children were unable to provide explanations to their peer's writing performance. Most of the time, they preferred short answers (*e.g.* Yes/No) compared to explanations, as shown in Fig. 3. We believe that this may be because of their young age (4-6 years old), and they might not have knowledge

<sup>4</sup><https://tla.mpi.nl/tools/tla-tools/elan/>

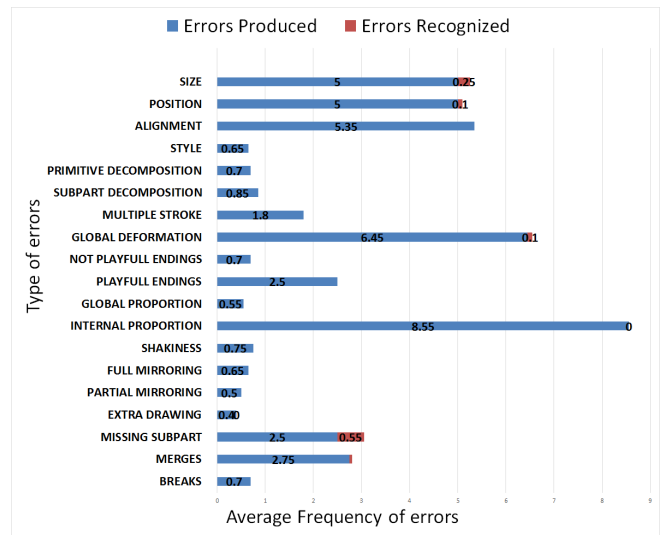


Figure 2: Types of handwriting errors produced and recognised by children

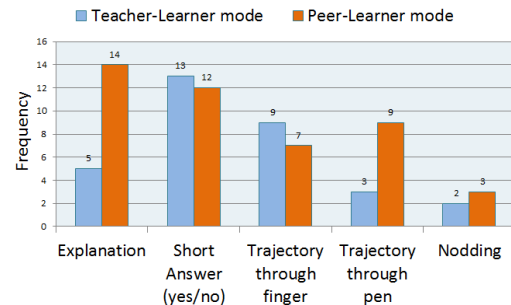


Figure 3: Communication modalities used by children for providing corrective feedback

concerning the shapes of letters. The second most-used communication modality was gestures - demonstrating the correct trajectory of a letter through a finger or a pen (see Fig. 3).

Regarding the difficulties in letter writing, we observed that for most of the children, the letters  $\alpha$  and  $\pi$  were difficult to write compared to the other letters. This difficulty might have arisen due to unfamiliarity of these letters. In general, children were able to understand both modes, but from the experimenter's point of view, the children showed more excitement in the "teacher-learner mode" compared to the "peer-learner mode", as they seemed excited to act as a 'teacher'. The built taxonomy was further revised in the second study and detailed in section *Study II* (see Fig. 4). The frequencies of the errors produced and recognised by the children are shown in Fig. 2.

## STUDY II

The results of the first study did not present sufficient evidence to differentiate the effectiveness of the two interaction modes. Additionally, we observed that the children were immature in providing corrective feedback on their peers' writing performance due to their young age. Thus, we conducted a second study that consisted of a between-subject design with 40 children in the age group of 6-8 years. We replicated the protocol of the first study but replaced the human facilitator



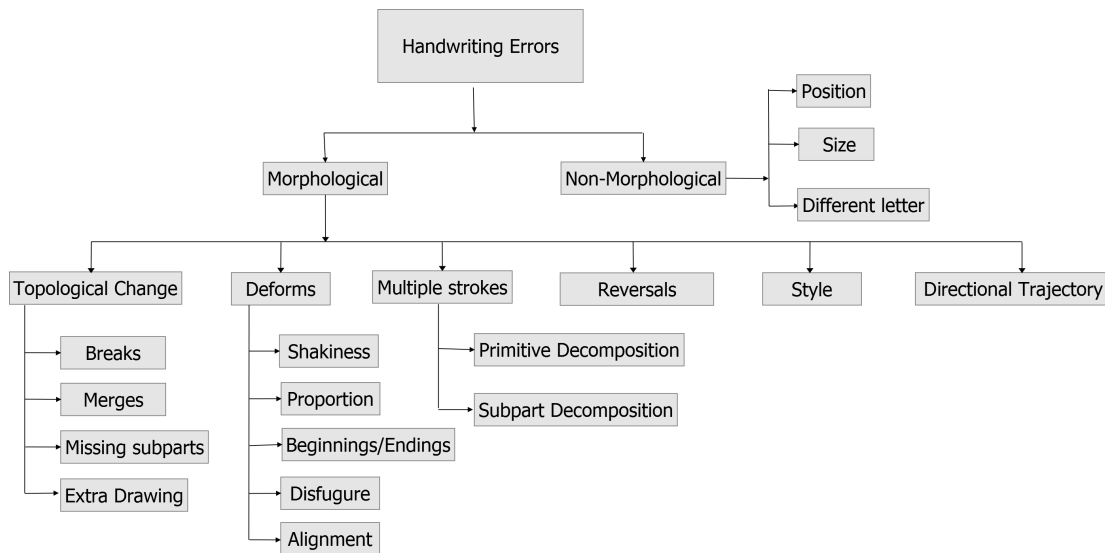


Figure 4: Taxonomy of errors



Figure 5: Children interacting with a robot under the PL/PT conditions

with a robot facilitator. The material in the study included a tablet and a pen for each child, installed with the writing app (used in *Study 1*). For the pre-post test, paper sheets with printed letters (j, D, K, y, W, t,  $\alpha$ , and  $\pi$ ) were used, while for the writing activity, 4 cards were created, and each card had a letter or word written on it: h, Moon, Ice-cream and Rainbow. In addition, we also used cameras, microphones and the Aldebaran's Nao robot (torso part)<sup>5</sup>.

### Goals

The overall goal of this study was to further investigate the writing errors of children in the age group of 6-8 years, revise the built taxonomy, and explore the PL and PT interaction modes, more precisely, in terms of the assessment of children's behavior with their peers while performing a collaborative writing activity. The research questions of the study include the following:

- 1) Would children 6-8 years of age perform similar writing errors as children 4-6 years of age?
- 2) Would children in the age group of 6-8 years be able to provide corrective feedback on the written performance of their peers?

- 3) How do the two interaction modes differ in terms of corrective feedback? Which type of corrective feedback (short feedback or extended feedback) would children provide?

### Methodology

To collect the handwriting samples of children and to explore the impact of introducing a robotic facilitator, the study was performed under two conditions: 1) a pair of children with a robot facilitator, performing a collaborative writing activity based on the PT method, and 2) a pair of children with a robot facilitator, performing a collaborative writing activity based on the PL method. The children were assigned the role of a teacher or a learner according to the learning method used under each condition. The role of the robot-facilitator was to provide instructions and interaction flow during the activity. A Wizard-of-Oz procedure was used to control the robot's behavior. A total of 16 children (8 pairs) participated in Condition 1, and 24 children participated in Condition 2. Each interaction session lasted approximately 10-15 minutes. The study followed these phases:

**Pre-test:** In this phase, a researcher guided two children in a study room and asked them to copy the letters onto their paper sheets.

**Interaction with the robot:** After completing the pre-test, the children were guided to a room where the robot facilitator was already placed. The researcher explained the study and left the room. Soon after, the interaction started with the introduction of the robot to the children. Then, the robot explained the features of the writing app and the writing activity. After that, the robot randomly assigned roles to the children according to the conditions (PT or PL). Under the PT condition, the robot asked the tutor-child to pick a card and show it to the learner-child so that he/she could write the letter or word on the tablet application. Then, the facilitator prompted the teacher-child to provide corrective feedback on the learner-child's written performance (see Fig. 5). On the other hand, under the PL

<sup>5</sup>Aldebaran robotics: <https://www.aldebaran.com/en>

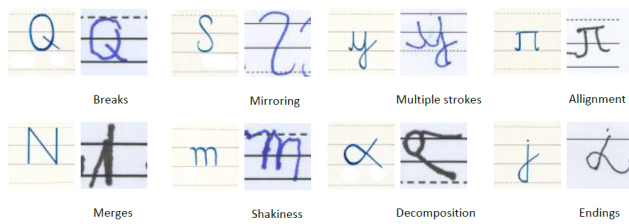


Figure 6: Children's handwritten letters corresponding to the errors. The correct samples of letters are shown on the left side of each letter.

condition, no roles were assigned to the children, and both acted as a teacher and as a learner. The robot alternately asked one of the children to pick a card. Both were instructed to write a letter/word on the tablet and prompted to provide feedback on each other's performance (see Fig. 5). Again, the process was repeated for all the letters/words. In the last phase of interaction, the robot thanked the children for their time.

**Post-test:** After finishing the interaction with the robot, the children were guided to the same room where they had performed the pre-test. The children then repeated the pre-test, which also served as post-test.

### Analysis

We collected the data from the tablets and paper sheets and analysed 60% of the data with two independent coders. Furthermore, we performed video analysis of all the groups by annotating the verbal communication between the children and the robot facilitator. We again used the ELAN tool with two independent coders. The agreement value of Cohen's kappa was .58 for the writing errors, which is considered as moderate agreement. The agreement value of Cohen's kappa was .92 for the verbal behaviors and .80 for the pre-/post-test scores. For further analysis, we used the non-parametric Exact Wilcoxon rank-sum (Mann-Whitney U) test, which also seems appropriate for the unequal sample size of the data.

### Results

#### Taxonomy of handwriting errors

After analysing the incorrect letters from the collected database of children 6-8 years of age, we revised the previous taxonomy to incorporate modifications based on the second study. As shown in Fig. 4, the error tree is classified into 17 types of handwriting errors. The classification of the tree is based on the English alphabet, including both uppercase and lowercase letters. However, in the first and second studies, two Greek letters were also included to check the expertise of the children, as they were unfamiliar with the letters. Moreover, the classification included different writing styles, for example, manuscript and cursive.

We classified the writing errors into two broad categories: morphological and non-morphological. These categories are further divided into subcategories (see Fig. 4).

**Morphology** is defined here as the study of the shapes of letters. The writing issues that show a change in the morphology of a letter belong to this category and are further divided into 6

subcategories: *topology changes*, *deforms*, *multiple strokes*, *reversals*, *directional trajectory*, and *style*.

**1- Topology change-** Topology is related to the properties of shapes that are preserved under continuous deformations, for example, bending and shrinking without breaking or merging. The writing issues that show a change in the topology of a letter belong to this category. This category contains four more subcategories explained below: <sup>6</sup>

- **Breaks-** Any discontinuity in the shape of a letter can be observed as a 'hole' in the letter. The writing errors that contain breaks or holes belong to this category (see Fig. 6)
- **Merges-** When a sub-part of a letter is glued to another sub-part of the same letter, it creates a merging error in the letter (see Fig. 6).
- **Missing Sub-parts-** Only those letters that have more than one stroke can have this error. A missing stroke in a letter creates this type of error.
- **Extra drawing-** This type of error refers to when an additional stroke is added (not necessarily touching the letter) to make a letter decorative or exaggerated. However, this is not considered to be a major error, as it generally adds artistic value to a letter.

**2- Deforms-** When a letter shows any deformation in the shape of the letter while maintaining the topology of the letter, the error belongs to this category. The deformations can be present in a sub-part of the letter (locally) or in all parts of the letter (globally). This category is further divided into 5 subcategories, discussed below:

- **Shakiness-** This type of error occurs when a letter shows a visible jerk in its trajectory. (see Fig.6).
- **Proportion-** This type of error occurs when a letter shows an asymmetry or disproportion in its shape, e.g. shrinking or stretching in any direction.
- **Beginnings/Endings**

This issue is when then beginning or ending is extended (in the shape of a curve or a straight line). This type of error is different from the *extra drawing* error, as it only exaggerates the extensions. This error is further divided into two additional categories, detailed below.

**Playful-** When a letter shows a decorative extension and the child who formed this letter is aware of the extension, then the error belongs to this category.

**Non-playful-**When a letter shows extensions (not decorative) and a child does not have knowledge of forming the extensions, then the error belongs to this category. However, it is difficult to identify this issue by looking at the final images of letters. It can be easily seen during the formation of a letter.

- **Distortion**

This is a general category of errors that do not follow any specific patterns. This type may include, for example, wide, curvy, and snake-style distortions in the shapes of letter and sometimes leads to fully distorted shapes resulting in illegible letters.

<sup>6</sup><https://en.wikipedia.org/wiki/Topology>

- **Alignment**

This refers to the misalignment or rotation of a letter through some angle (see Fig.6).

**3- Multiple strokes-** When a letter is formed by decomposing it into parts, this type of error is created (see Fig.6). If the parts are well joined in the final image of the letter, then it is difficult to visualise this error; however, during the formation of a letter, this issue can be easily observed. In addition, the error may or may not change the topology of the letter and can be present locally or globally. This category contains two subcategories, detailed below:

- **Primitive Decomposition-** This type of issue is when a letter is formed by decomposing it into basic shapes, for example, circles and horizontal/vertical lines (see Fig. 6). In the first study, we observed this error mostly in beginners, and it may occur due to the lack of knowledge about shapes. For example, when writing the letter *alpha*, a few children decomposed it into a circle and straight lines, as the letter was unfamiliar to them (see Fig.6).
- **Sub-part Decomposition-** This issue refers to when a letter is formed by decomposing a single stroke into several strokes. This is different from the above error, as in this case, the decomposed strokes do not resemble basic shapes such as a circle. Here, a circle can possibly be made with several small strokes.

**4- Reversals-** This error refers to when a subpart or a full letter is reversed in any of the four directions, left, right, top, or bottom; then, a reversed image is created. This does not affect the topology of the letter.

**5- Direction of trajectory-** This writing error can only be observed during the formation of a letter, as it concerns the direction of the trajectory of the letter. This issue does not produce a change in the topology of the letter and can be present locally or globally.

**6- Style-** This issue refers to when a letter is formed with a different style (cursive or manuscript) instead of the required style. However, this is a minor issue and can be present locally or globally. The topology may or may not be changed.

**Non-morphological** errors do not produce a change in the shape of a letter and therefore are not related to the morphology of a letter. In addition, this error does not affect the topology of the letter. This category is divided into three additional subcategories, discussed below.

**1- Position-** This issue refers to the correct placement of a letter with respect to base/reference lines present on a writing base where the letter is drawn.

**2- Size-** As the name suggests, this issue is related to the size of a letter with respect to reference lines or other letters present on a writing base where the letter is drawn.

**3- Different letter-** This error is related to the misinterpretation of a letter. For example, in both studies, we found two letters,  $\alpha$  and  $\pi$ , were misinterpreted as the letters “a” and “r” or “u”, respectively. Initially, we did not include this issue in the tree because we thought that this error might occur due to the introduction of unknown letters to the children. However, in the second study, we found that two children misinterpreted the letters “W” and “j” as ‘U’ and ‘G’, respectively, which

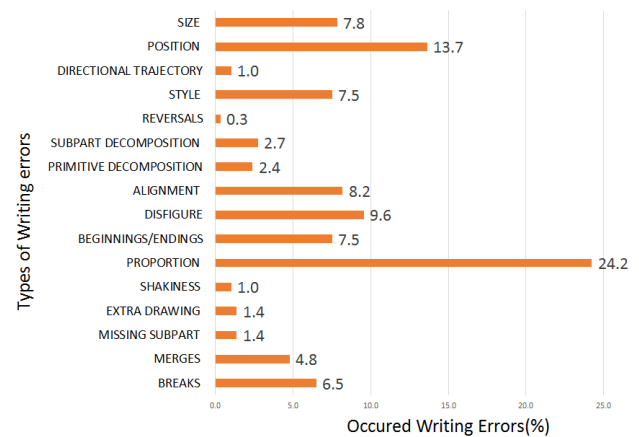


Figure 7: Handwriting errors that occurred (%)

shows that this error may occur, although not often.

The above writing issues were explored by analysing the letter images as well as by observing the formation of letters during both studies. The classification is based on considering the change in shape of the letters and patterns of the writing errors. We revised the taxonomy in the second study and modified it by adding and merging some writing issues. For example, the writing errors *decomposition* and *multiple strokes* were merged, and the error *different letter* was added in the revisions. Although we observed that a few writing errors, such as shakiness, missing subpart and extra drawing, seemed to be less frequent in the second study (see Fig. 7), other errors remained prevalent in both studies. The frequencies of the errors in the second study are shown in Fig. 7, and a few samples of the handwritten letters from the second study are presented in Fig. 8.

**Exploration on modes of interaction & Children’s assessment on their peers**

The two modes of interaction, the PL and PT learning methods, were explored in terms of three variables: 1) self-disclosure, 2) corrective feedback, and 3) learning gains. Regarding self-disclosure, all the tutor-children and learner-children under the PT condition gave significantly more extended self-disclosures to the robot in comparison to the children under the PL condition [11]. Regarding the corrective feedback given by the children, the result of a Man-Whitney U test suggested that all learner-children under the PT condition gave significantly more extended feedback to their peers (Mdn = 5.0) compared to the children under the PL condition (Mdn = 0.0),  $U = 57$ ,  $p < 0.05$ ,  $r = .43$ , as depicted in Fig. 9(a) [11]. On the other hand, the results showed that all the tutor-children under the PT condition gave significantly more extended corrective feedback to their peers (Mdn = 2.5) in comparison to the children under the PL condition (Mdn = 0.0),  $U = 21$ ,  $p < .001$ ,  $r = .73$ , as observed in Fig. 9(b) [11]. The results of the children’s learning gains under the PT condition suggest a significant improvement in the children’s learning gains in the post-test compared to the pre-test. However, no significance differences in pre-/post-test scores were found under the PL condition [11]. Overall, the PT method found to be more effective in all three regards.

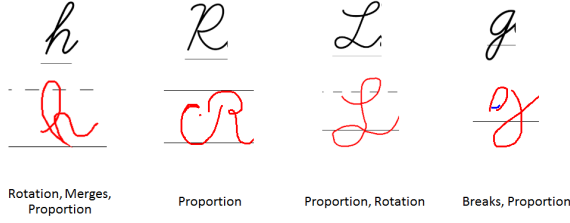


Figure 8: Children's handwriting issues

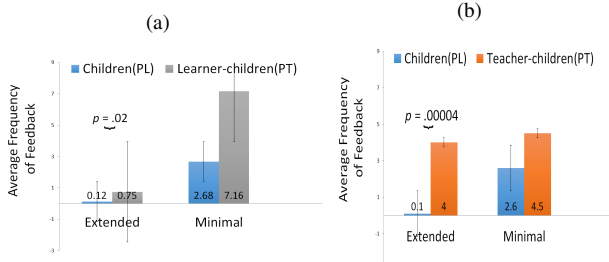


Figure 9: Results of the corrective feedback between (a) learner-children under the PT condition and all children under the PL condition and between (b) tutor-children under the PT condition and all children under the PL condition.

#### IMPLEMENTATION OF WRITING ERRORS FOR A ROBOT

To test some writing errors from the developed taxonomy, we implemented a set of errors in a system that provides an educational scenario for acquiring the writing skills. The scenario consists of a child, whereby s/he corrects the writing errors of a robot, which acts as a learner. The scenario attempts to demonstrate how children comprehend these writing errors, which in reality depicts their own writing errors.

Generating the deformed letters for the learner-robot and correcting a letter (from the child's side) was a key point in the scenario. Instead of generating the letters with random handwriting errors for the learner-robot, we selected three common errors, *Proportion*, *Breaks* and *Alignment*, from the built taxonomy of handwriting errors.

The choice of using these handwriting issues was motivated by two reasons: their prevalence in children's handwriting [16, 17] and because we also observed in our past studies that the *Proportion* and *Alignment* issues were the most commonly occurring writing errors. We chose only three errors to keep the scenario simple and to prevent stress on a child's cognitive load, particularly, by not adding more than three errors in the same scenario. To handle the *proportion* and *breaks* writing errors, we used an algorithm proposed in [38], which combines the output of multiple algorithms in an inverse control problem. Additionally, the algorithm also incorporates human-movement-inspired features and also generalises the synthesis to generate poor or good written samples of a letter. We used the algorithm for two primary reasons: First, the algorithm provides the capability to synthesise and learn the multiple-mode motion trajectories through the integration of rapid representation and extraction. Second, we needed to generate deformed letters for the learner-robot that resemble children's handwriting.

The input of the algorithm is a planar handwriting dataset  $\mathcal{D}$  (see Algorithm 1). To generate similar instances of a letter present in the dataset, the Gaussian statistics of the trajectories are extracted. The parameters received from the statistics are further converted into a curvature-based representation with a nonlinear feature embedding. The mean  $\mu_f^m = \{A_j, z_0^j, \mu_j, \sigma_j, \alpha_s^j, \alpha_e^j\}_{j=1:N}$  is obtained by an RXZero process, which effectively solves a nonlinear optimization to match the trajectories through Equations (1) and (2). The covariances are transformed through local linear projections, with  $J(\cdot)$  denoting the Jacobian of the reconstruction. It is easy to find that some of the extracted parameters are correlated to the motion path in an explicit manner.

$$|v(z)| = \sum_{j=1}^N \frac{A_j}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(\ln(z-z_0^j) - \mu_j)^2}{2\sigma_j^2}\right) \quad (1)$$

$$\phi_j(z) = \alpha_s^j + \frac{\alpha_e^j - \alpha_s^j}{2} \left(1 + \operatorname{erf}\left(\frac{\ln(z-z_0^j) - \mu_j}{2\sigma_j}\right)\right) \quad (2)$$

---

#### Algorithm 1 Learning with curvature-based features for modelling handwriting motion

---

**Require:**  $\mathcal{D} = \{\zeta_i\}, M$

**Ensure:**  $\mathcal{D}_{k=1:K_m}^m, \hat{\theta}_k^m = \{\mu_f^m, \Sigma_f^m\}, m = 1, \dots, M$

$\mathcal{D}_{k=1:K_m}^m \leftarrow \text{Partitioning}(\mathcal{D}, M)$   $\triangleright$  Build locally similar dataset  $\mathcal{D}$  through partitioning

**for all**  $m$  in  $1:M$  **do**

$$\mu_\zeta^m, \Sigma_\zeta^m \leftarrow \operatorname{argmax}_{\theta} \sum_{i=1}^{|\mathcal{D}_k^m|} \log \mathcal{N}(\zeta_i^* | \theta)$$

$$\mu_f^m \leftarrow \text{RXZERO}(\mu_\zeta^m)$$

$$\Sigma_f^m \leftarrow J(\mu_\zeta^m)^T \Sigma_\zeta^m J(\mu_\zeta^m)$$

**end for**

---



Figure 10: Letter samples generated from the algorithm showing the modulation in the components producing the *proportion* and *break* handwriting errors.

For instance, in Equation 1,  $A_j$  regulates the velocity magnitude so that the proportion of specific letter components can be modulated without influencing the other components. Equation 2 shows that  $\alpha_s^j$  and  $\alpha_e^j$  control the angular position. Hence, the two parameters allow for the stroke orientation or the curvature to be adjusted; see Fig. 10. For the *alignment* issue, we developed an algorithm to produce the rotation effect with different angles.

#### VALIDATION

In our child-robot learning scenario of the final study, we were interested in exploring two different aspects: 1) how children would perceive the writing abilities, specifically, the writing errors of the robot, and 2) how children would perceive the



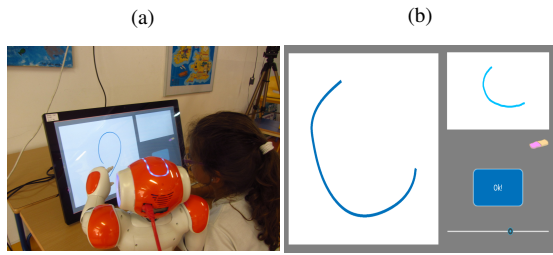


Figure 11: (a) A child and a robot performing the writing activity; (b) Interactive screen

behavior of the robotic agent, as this may affect the child-robot learning interaction. Thus, we tested the above system by conducting a 4-week study (1 interaction per week) with 24 children in the age group of 7-8 years at a school. Considering the results of the second study, which revealed the richness of the PT learning method, the current system relies on the PT learning method. The study consists of between-subject design and has two conditions, therein demonstrating the different competencies of the learner-robot: *Learning* and *Non-Learning*. A total of 13 children ( $M=8$ ;  $SD= 0.784$ ; 8 male and 5 female) participated under the *learning condition*, while 11 children ( $M=7.92$ ;  $SD=.828$  years old; 5 male and 8 female) participated under the *non-learning condition*, and each interaction lasted approximately 12-15 minutes. The scenario involved a tutor-child correcting the handwriting errors of a learner-robot (see Fig. 11(a)). As shown in 11(b), the screen shows interactive features, such as a small white box (right side), and a slider. The slider was connected to the algorithm (Algorithm 1), which has the ability change the shape of a letter by moving the slider. The tutor-child provided corrections in two ways - by moving the slider and by providing a demonstration of the letter in the small box.

Under the learning condition, the robot showed a learning progression after each interaction, whereas under the non-learning condition, the robot showed a consistent performance (it did not learn) throughout the study. To provide spontaneous and fluid interaction between the child and robot, an autonomous social behavior for the robot was developed. Following the interaction, the experimenter asked self-response questions to the children based on a 5-point Likert scale and partially inspired by the Godspeed questionnaire [1]. The questions were asked at three time intervals - immediately after the second, third and fourth interactions with the robot.

The overall goals of the study concern the testing of the writing errors considering child-robot interactions and exploring children's perception of the writing abilities and social behavior of the *Learning* vs. *Non-Learning* robot through child self-response questionnaires. Moreover, the study explores the following research questions: 1- Can children differentiate the learning abilities of the robot between the two conditions? The learning abilities are related to the robot's writing ability and overall performance. 2- Would the two different competencies (learning and non-learning) of the robot affect the children's perceived likability and friendliness towards it?

We collected the data from the questionnaires and analysed them using the non-parametric Mann-Whitney U test. The pre-

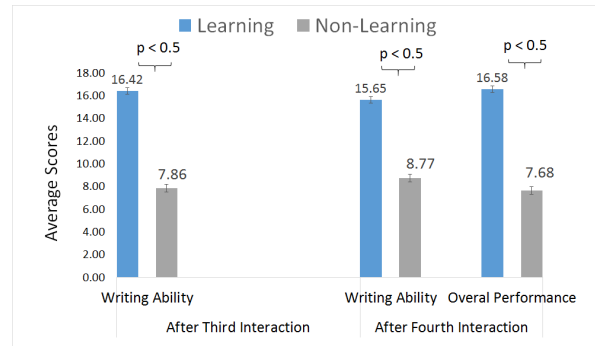


Figure 12: Results of the robot's learning scores given by the children: Writing ability and overall performance scores after the third and fourth interactions.

liminary results related to the children's perception of the robot are presented in the paper. Questions related to the writing ability and the overall performance of the robot showed significant differences over the four-week interactions. Despite showing no differences in the first two interactions, after the third interaction, children under the learning condition (mean rank = 16.42) gave significantly higher *writing ability scores* to the robot compared to the non-learning condition (mean rank = 7.86),  $U = 20.5$ ,  $z = 86.5$ ,  $p = .002$  (see Fig. 12). After the fourth interaction, children under the Learning condition (mean rank = 16.58) gave significantly higher *overall performance* scores to the robot compared to the Non-learning condition (mean rank = 7.68),  $U = 18.5$ ,  $z = -3.366$ ,  $p = .001$ . Additionally, for the writing ability, under the Learning condition (mean rank = 15.65), they gave higher scores to the robot compared to the Non-learning condition (mean rank = 8.77),  $U = 30.50$ ,  $z = -2.67$ ,  $p = .015$  (see Fig. 12). The result suggests that children indeed paid attention to the robot's writing errors and hence were able to notice the difference in the robot's learning abilities over time.

Regarding children's perceived social behavior, we observed that after the last interaction, 100% of the children under the learning condition and 92% under the non-learning condition gave high scores for both likeness and friendliness questions. Furthermore, concerning the perceived intelligence of the robot, 100% of the children under the learning condition and 67% under the non-learning condition gave high scores to the robot, even when being aware of the robot's lack of learning abilities. These results suggest that the children's perceived social behavior was not affected by the robot's learning and non-learning competencies.

## CONCLUSION

We believe that this article presents the following important contributions: the proposition of a taxonomy of writing errors, based on the handwritten letters of children 4-8 years of age, that may be useful for educational purposes; the exploration modes of interactions between children and a robot during the first two studies; the test of a set of writing errors from the taxonomy in an educational scenario whereby a child corrects the writing errors produced by a learner-robot; and finally the exploration of children's perception towards a robot with learning vs. non-learning competencies.

The presented taxonomy of the writing errors provides a deep understanding of common handwriting issues for letters of the English alphabet (Uppercase and Lowercase). We believe that the taxonomy extends the existing literature, as it contains a few writing issues (e.g. *multiple strokes and breaks*) that have not been explored previously. Moreover, the classification in the taxonomy may provide a new understanding of writing errors. It is worth noting that the writing errors defined in the taxonomy may benefit both teachers and students. Generally, if a student struggles in forming a letter, s/he practices it by writing it several times, which may involve the correction of random writing errors in the same letter. However, one of the key advantages of using the taxonomy is in helping a child with one particular writing error at a time. In addition, the proposed taxonomy can be used to evaluate children's handwriting skills in a more structured way. To improve handwriting, researchers may develop an educational scenario dedicated to a particular writing issue or similar writing issues present in the taxonomy. However, incorporating all the handwriting issues for one educational activity seems to be challenging, as writing errors differ with respect to the levels of handwriting and motor-cognitive skills. In addition, this would also increase the cognitive load on a child.

The implementation of the three writing errors in the third study indicated children's attention towards the robot's writing errors as well as improvements. These results suggest that correcting the errors (which resembles children's own writing errors) may allow children to pay attention to their own writing errors. Finally, regarding child-robot interactions, the work presented here investigates two modes of interactions and two different roles of a robot in different learning scenarios (the second and third studies), which may be helpful for the human-robot interaction community. In contrast to the second study, the third study provides an active role to the robot, as a learner, which directly involves the writing process and affects children's perception towards the robot's writing abilities.

## FUTURE WORK

Future work includes further revisions of the current taxonomy by conducting more child-robot studies. We also plan to prepare a database of children's handwritten words, therein using the English alphabet and collected through future child-robot studies, while exploring the writing issues present in the words. It would be meaningful to compare the two taxonomies, the 'letter' and 'word' taxonomies, to see whether some of the writing errors from the 'letter' taxonomy would disappear.

## ACKNOWLEDGMENTS

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013 and through project AMIGOS (PTDC/EEISII/7174/2014). We show our gratitude to all the schools involved in the studies.

## REFERENCES

1. Christoph Bartneck, Elizabeth Croft, and Dana Kulic. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics* 1, 1 (2009), 71–81. DOI: <http://dx.doi.org/10.1007/s12369-008-0001-3>
2. Mary Benbow, Barbara E Hanft, and Dottie Marsh. 1992. *Handwriting in the classroom: Improving written communication*. American Occupational Therapy Association.
3. Tanya N. Beran, Alejandro Ramirez-Serrano, Roman Kuzyk, Meghann Fior, and Sarah Nugent. 2011. Understanding How Children Understand Robots: Perceived Animism in Child-robot Interaction. *Int. J. Hum.-Comput. Stud.* 69, 7-8 (July 2011), 539–550. DOI: <http://dx.doi.org/10.1016/j.ijhcs.2011.04.003>
4. Virginia W Berninger and Dagmar Amtmann. 2003. Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. (2003).
5. Virginia W Berninger and Steve Graham. 1998. Language by hand: A synthesis of a decade of research on handwriting. *Handwriting Review* 12, 1 (1998), 11–25.
6. K Berry. 1989. The VMI: Developmental Test of Visual-motor Integration (Cleveland. (1989).
7. Gautam Biswas, Krittaya Leelawong, Daniel Schwartz, Nancy Vye, and The Teachable Agents Group at Vanderbilt. 2005. Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence* 19, 3-4 (2005), 363–392.
8. Mary-Ann Bonney. 1992. Understanding and assessing handwriting difficulty: Perspectives from the literature. *Australian Occupational Therapy Journal* 39, 3 (1992), 7–15.
9. David Boud, Ruth Cohen, and Jane Sampson. 2001. Peer learning and assessment. *Peer Learning in Higher Education: Learning from & with Each Other* (2001), 67.
10. Ginevra Castellano, Ana Paiva, Arvid Kappas, Ruth Aylett, Helen Hastie, Wolmet Barendregt, Fernando Nabais, and Susan Bull. 2013. Towards Empathic Virtual and Robotic Tutors. In *Artificial Intelligence in Education*, H.Chad Lane, Kalina Yacef, Jack Mostow, and Philip Pavlik (Eds.). Lecture Notes in Computer Science, Vol. 7926. Springer Berlin Heidelberg, 733–736.
11. Shruti Chandra, P. A.Oliveira, Séverin Lemaignan, Pedro Sequeira, Ana Paiva, and Pierre Dillenbourg. 2016. Children's peer assessment and self-disclosure in the presence of an educational robot. In *Robot and Human Interactive Comm.(RO-MAN), 2016 25th IEEE Int. Symposium on*. IEEE, 539–544.
12. Clinton I Chase. 1986. Essay test scoring: Interaction of relevant variables. *Journal of Educational Measurement* 23, 1 (1986), 33–41.
13. Carol A Christensen. 2005. The role of orthographic–Motor integration in the production of creative and well-structured written text for students in secondary school. *Educational Psychology* 25, 5 (2005), 441–453.

14. CE Exner. 1989. Development of hand functions. *Occupational therapy for children* (1989), 235–259.
15. Katya P Feder and Annette Majnemer. 2007. Handwriting development, competency, and intervention. *Developmental Medicine & Child Neurology* 49, 4 (2007), 312–317.
16. Steve Graham, Karen R Harris, Linda Mason, Barbara F. Chorzempa, Susan Moran, and Bruce Saddler. 2008. How do primary grade teachers teach handwriting? A national survey. *Reading and Writing* 21, 1-2 (2008), 49–69.
17. Steve Graham, Naomi Weintraub, and Virginia Berninger. 2001. Which manuscript letters do primary grade children write legibly? *Journal of Educational Psychology* 93, 3 (2001), 488.
18. Deanna Hood, Séverin Lemaignan, and Pierre Dillenbourg. 2015. When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting. In *Proceedings of the Tenth Annual ACM/IEEE Int. Conf. on Human-Robot Interaction*. ACM, 83–90.
19. Alexis D Jacq, Séverin Lemaignan, Fernando Garcia, Pierre Dillenbourg, and Ana Paiva. 2016. Building successful long child-robot interactions in a learning context. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 239–246.
20. Jeffrey Johnson. 2003. Children, robotics, and education. *Artificial Life and Robotics* 7, 1–2 (2003), 16–21.
21. Peter H Kahn Jr, Takayuki Kanda, Hiroshi Ishiguro, Nathan G Freier, Rachel L Severson, Brian T Gill, Jolina H Ruckert, and Solace Shen. 2012. “Robovie, you’ll have to go into the closet now”: Children’s social and moral relationships with a humanoid robot. *Developmental psychology* 48, 2 (2012), 303.
22. Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. 2004. Interactive Robots As Social Partners and Peer Tutors for Children: A Field Trial. *Human-Computer Interaction* 19, 1 (June 2004), 61–84.
23. Takayuki Kanda, Rumi Sato, Naoki Saiwaki, and Hiroshi Ishiguro. 2007. A two-month field trial in an elementary school for long-term human–robot interaction. *Robotics, IEEE Transactions on* 23, 5 (2007), 962–971.
24. Ragnheidur Karlsdottir and Thorarinn Stefansson. 2002. Problems in developing functional handwriting. *Perceptual and motor skills* 94, 2 (2002), 623–662.
25. Mike Keppell, Eliza Au, Ada Ma, and Christine Chan. 2006. Peer learning and learning-oriented assessment in technology-enhanced environments. *Assessment & Evaluation in Higher Education* 31, 4 (2006), 453–464.
26. Judith I Laszlo and Phillip J Bairstow. 1984. Handwriting: Difficulties and possible solutions. *School Psychology International* 5, 4 (1984), 207–213.
27. Annlaug Flem Maeland. 1992. Handwriting and perceptual-motor skills in clumsy, dysgraphic, and “normal” children. *Perceptual and motor skills* 75, 3 suppl (1992), 1207–1217.
28. Theresa Malloy-Miller, Helene Polatajko, and Bev Anstett. 1995. Handwriting error patterns of children with mild motor difficulties. *Canadian Journal of Occupational Therapy* 62, 5 (1995), 258–267.
29. Shizuko Matsuzoe and Fumihide Tanaka. 2012. How smartly should robots behave?: Comparative investigation on the learning ability of a care-receiving robot. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Comm.* IEEE, 339–344.
30. Jane Medwell and David Wray. 2008. Handwriting—a forgotten language skill? *Language and Education* 22, 1 (2008), 34–47.  
<http://www.tandfonline.com/doi/abs/10.2167/1e722.0>
31. Adrian D Sandler, Thomas E Watson, Marianna Footo, Melvin D Levine, William L Coleman, and Stephen R Hooper. 1992. Neurodevelopmental study of writing disorders in middle childhood. *Journal of Developmental & Behavioral Pediatrics* 13, 1 (1992), 17–23.
32. Rosemary Sassoon. 1990. *Handwriting: A new perspective*. Stanley Thornes.
33. ML Simner. 1991. Estimating a child’s learning potential from form errors in a child’s printing. *Development of graphics skills: Research, perspectives, and educational implications* (1991), 205–222.
34. Marvin L Simner. 1982. Printing errors in kindergarten and the prediction of academic performance. *Journal of Learning Disabilities* 15, 3 (1982), 155–159.
35. SR Stewart and CS Simon. 1985. Development of written language proficiency: Methods for teaching text structure. *Communication skills and classroom success* (1985), 341–361.
36. Carol O Sweedler-Brown. 1992. The Effect of Training on the Appearance Bias of Holistic Essay Graders. *Journal of Research and Development in Education* 26, 1 (1992), 24–29.
37. Fumihide Tanaka and Shizuko Matsuzoe. 2012. Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction* 1, 1 (2012).
38. H. Yin, P. A. Olivera, F. S. Melo, A. Billard, and A. Paiva. 2016. Synthesizing Robotic Handwriting Motion by Learning from Human Demonstrations. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*. (2016).