# A game-based approach for evaluating and customizing handwriting training using an autonomous social robot

Daniel Carnieto Tozadore[1,2], Chenyang Wang[1], Giorgia Marchesi[3], Barbara Bruno[1], Pierre Dillenbourg[1]

***Abstract*— Handwriting learning is a long and complex process that takes about ten years to be fully mastered. Nearly one-third of all children aged 4-12 experiences handwriting difficulties and, sadly, most of them are left to fight them on their own, due to the scarcity of tools for the detection and remediation of such difficulties. Building on state-of-the-art digital solutions for automated handwriting assessment and the training of specific handwriting-related skills, in this article we discuss requirements, rationale, and architecture of a system for handwriting training, which relies on a social robot as a mediator agent, offering personalized training and suggestions. The system is envisioned to operate autonomously and to support long-term interactions via personalization. Preliminary validation of the system in an experiment with 31 children showed its potential not only for autonomously guiding handwriting training sessions, but also for its inclusion in the teachers' practice.**

## I. Introduction

How many people do you know with beautiful handwriting? If the number is relatively low, don't be surprised: around one third of children between 4 to 12 years old presents handwriting difficulties [1], which are not always properly addressed and corrected. Since such difficulties can cause tremendous and long-lasting damage to a child, to the point of inspiring school avoidance and low self-esteem, early detection and remediation are key for effectively overcoming them [2]. A major hurdle towards this goal is the *vicious circle* that handwriting difficulties create: children with handwriting difficulties find hard to write and obtain lesser results than their peers. Hence, they avoid writing as much as possible, which results in lagging further behind their peers and an even stronger desire to avoid writing altogether [3]. For this reason, remediation activities, which necessarily have to involve writing sooner or later, are typically met with significant resistance.

In an attempt to overcome this resistance, a number of approaches have been recently proposed which rely on digital devices such as tablets [4] and robots [5]. On the one hand, such approaches rely on the *novelty effect* and *gamification* to counterbalance a child's reluctance to engage in handwriting practice. Recently, the rich sensory information made available by digital devices allows for monitoring the child's progress and the penalization of the exercises, at levels previously unreachable and with inspiring results [6].

This project has received funding from the Swiss National Science Foundation through the National Centre of Competence in Research (NCCR) Robotics and through project iReCHeCk (FNS 200021E_189475/1).

[1]Computer-Human Interaction in Learning and Instruction (CHILI) Lab, EPFL, Lausanne, Switzerland. [2] MOBOTS Group, EPFL, Lausanne, Switzerland. [3] Dept. DIBRIS, University of Genova, Italy. Email: daniel.tozadore@epfl.ch.

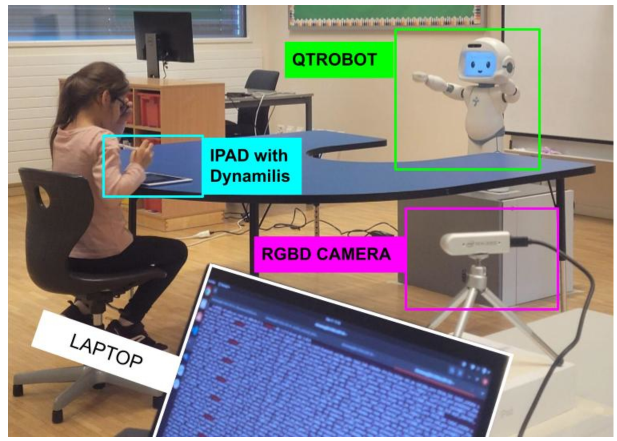


Fig. 1. The iReCHeCk hardware setup for a handwriting training session.

However, to date, the assessment of handwriting quality (and thus, the detection of handwriting difficulties) is still mostly done by experts through visual inspection of short handwritten sections, a process which is error-prone, slow and costly [7]. Above all, however, such an assessment method completely neglects the kinematics and dynamics of handwriting, such as one's speed, the smoothness of the motion, the pressure applied on the paper via the pencil, or its inclination. Recent works aiming at developing an automated handwriting assessment based on the real-time information provided by digital tablets and pencils have not only shown remarkable potentials [3], [7], but also highlighted the importance of correctly identifying the causes for one's difficulty and appropriately tailor remediation activities.

In this article, we outline and preliminarly validate the architecture of a system for handwriting training, which relies on a digital tablet for the automated, real-time assessment of handwriting quality and areas of improvements, and on a social robot for the automatization of the training sessions and the interaction with the child, as shown in Figure 1.

To test the system, we run an experiment in a school-environment where 31 children performed a single individual session guided by the robot. The session included games and exercises of handwriting on an iPad. The experiment suggests that the system seems robust and engaging enough for use in a school setting, by children. Additionally, although it is known that handwriting training is a long process, the real-time handwriting assessment methods afforded us to explore the existence of immediate effects of handwriting exercises on handwriting quality, an impracticable analysis prior to the development of automated, real-time handwriting assessment tools.

This paper is organized as follow. In Section II, related works that motivate and sustain this research are presented. Section III details the architecture of our proposal. In Section IV, the experiment is described and its results are reported in Section V. Conclusions follow.

## II. Related Work

The technical challenges in this proposal rely on two steps of detecting and remediating handwriting difficulties while the social aspects are about keep users motivated taking the outcomes as inputs.

Attempts at developing objective methods to assess the legibility (hence, the quality) of handwriting started with Thorndike, in 1910 [8] and can be divided into *holistic* approaches, which evaluate the quality of a handwritten text by comparing it with reference samples [8]–[10], and *atomistic* approaches, which evaluate the quality of a handwritten text on the basis of a set of pre-defined criteria, which are then summed up to produce an overall quality score. As an example, the *Concise Evaluation Scale for Children's Handwriting* (BHK) test, one of the primary references for the diagnosis of dysgraphia in Latin alphabet-based languages [11], follows the second approach.

Education and special education have proven to be a successful application field for social robots [12], which have therein been used for purposes as diverse as second-language learning [13], reading practice [14], or honing of computational thinking skills [15]. While in most of such applications the robot acts as a tutor and/or mediator between the child and a learning activity typically running on a tablet [12], a few exceptions envision the robot to be the learner and the child to be the teacher, in accordance with the *learning by teaching* paradigm [16]. Among them, the *CoWriter* project uses a Nao robot to help children practice handwriting [5].

In 2018/2019, a 10-years old boy diagnosed with a complex Neuro-Developmental Disorder combining phonological disorder, attention deficit/hyperactivity disorder, dyslexia, and developmental coordination disorder with severe dysgraphia, used *CoWriter* during his occupational therapy for 20 consecutive weekly sessions [6]. Results showed that his motivation was restored, avoidance behaviors disappeared both during sessions and at school, handwriting quality and posture improved dramatically. Most remarkably, the boy had had 2 years of specific support in school and professional speech and motor remediation prior to participation in this pilot, with no visible improvement.

While the longitudinal study highlighted the tremendous potential of robot-enhanced solutions for handwriting training, it also revealed a number of shortcomings that need to be addressed for a successful use of the system beyond controlled pilot studies: (1) the *CoWriter* setup used with the child was fully tele-operated, meaning that the boy's therapy sessions involved a robot, an occupational therapist managing the session and another domain expert (a psychiatrist) tele-operating the robot. Endowing the robot with *autonomy* is crucial to make the system less expensive, less complex, and thus more likely to be used at large in schools; (2) the novelty effect of the robot quickly wore off along the 20 sessions, revealing the importance of *penalization* and *adaptation* to sustain the child's engagement with the activity and keep up with his/her progress. Endowing a robot with such capabilities becomes even more important in the case it is expected to autonomously lead the sessions, with only sporadic interventions from therapists/teachers.

In respect to decision-making based on user modeling for supporting users motivation and engagement, several works have been striking this issue from simple algorithms modeling to more complex ones. For instance, a design using a Behavior Tree Based approach [17] for support long-term social robot behaviors showed to provide smooth and intuitive transitions between the states, also implemented by ROS libraries. This proposal, despite not validated with users yet, suggests efficient awareness of the system about when to take a decision according to users' actions with a low computational cost solution. Algorithms that take more time to train their model and, consequently, present more computational costs are also used to address this issue. Machine learning algorithms with supervised domain adaptation (s-DA) to afford personalized models are examples of this kind of knowledge base for decision-making [18]. Evaluations of such methods on the effects of personalization on a long-term multimodal dataset showed that their outcomes outperformed non-personalized, individualized and generic model baselines in both individual sessions and also in the average of all sessions. Beyond acquiring knowledge of users' preferences based on previous experiences, as supervised methods do, it is also possible to generate models based on the users action during execution time. A well-established approach is to adapt the robot's behavior according to the users' short-term inputs aiming to increase their engagement and enjoyment. As an example, the study adapting the behavior of a robot acting as an entertainer and telling different types of jokes [19]. The exemplary adaptation process was done only by using the audience's vocal laughs and visual smiles of 24 participants. When compared to a baseline behavior (non-personalized one), the adaptive setup showed better results regarding user engagement and enjoyment.

In this proposal, therefore, we aim to combine the beforehand mentioned elements of handwriting quality measurement and remediation combined to personalized and adaptive instructions from a social robot to keep young students motivated in their handwriting practicing.

## III. Architecture and Components

Figures 1 and 2 summarize, respectively at the hardware and software levels, the key elements composing the proposed system for handwriting training. Figure 1 shows from a hardware perspective the system, which includes the social robot QTrobot (from Luxai[1]), an iPad with Apple Pencil, running the *Dynamilis* app, an external RGB-D camera and a computer hosting the main software modules. The iPad and

[1] https://luxai.com/

QTrobot are placed in front of the child, while the external RGB-D camera is placed on the side.

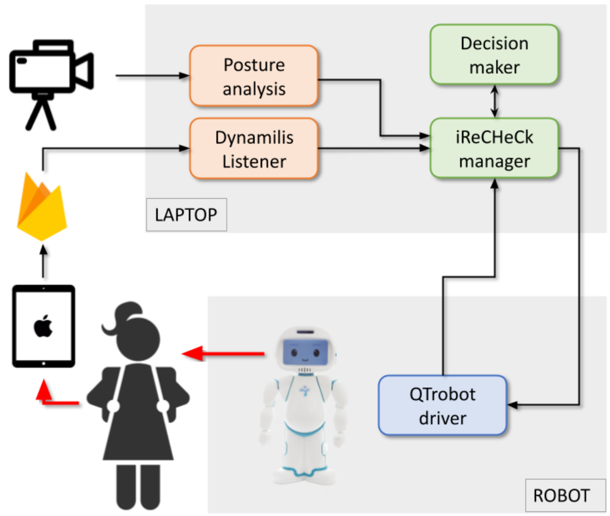


Fig. 2. iReCHeck system architecture. Sensing modules are denoted in red, planning ones in green and the action in blue. Red arrows denote non-digital information flows (the robot providing verbal suggestions and feedback to the child; the child interacting with *Dynamilis* activities via the Apple Pencil).

From a software perspective, the system is designed to be modular, allows for an easy adaptation to different configurations (e.g. with or without a human expert controlling the robot, with a rich or minimal external sensors set). Concretely, the modules are designed as ROS Noetic nodes and packages, written in Python 3[2]. The modules can be categorized in: (A) *sensing* modules, each of which interfaces to a specific type of perception device(s) and is responsible for the extraction of high-level information from it. As an example, the *Dynamilis listener* module is responsible for obtaining handwriting-related information from the *Dynamilis* app, while the *Posture analysis* module process posture-related information from the RGB-D camera; (B) *planning* modules, where the information provided by the different sensing modules is merged and used to decide what to do/suggest next; (C) *action* module, which interface to the device(s) responsible for direct interaction with the child (currently, QTrobot), sending commands for execution and also providing information about users' action, such as when a user appears by face detection provided off-the-shelf via the Nuitrack[3].

### A. Dynamilis app and the Handwriting Evaluation (HWE)

The *Dynamilis* app[4], running on the iPad and requiring the use of an Apple Pencil, includes an activity for run-time automated handwriting assessment and a number of games for handwriting training, targeting different sub-skills related to handwriting (e.g., modulation of the pressure applied on the tablet via the pencil, or speed control while following a trajectory). More precisely, the *Handwriting Evaluation* (HWE) activity is a two-steps activity, first asking the child to draw a cat and then to copy a short text shown at the top of the screen. The reason for the two steps is to mitigate the negative attitude towards writing and decouple handwriting difficulties from reading difficulties. Each evaluation produces a handwriting quality score (expressed in the [0,100] range) along the dimensions of *static*, *speed*, *tilt* and *pressure*, plus a *global* score, following the methodology outlined in [20].

### B. Posture analysis

At this point of our implementation, the *posture analysis* node is an ongoing effort targeting an autonomous posture-analysis algorithm specifically for handwriting activities. We envision the future functionality of this node by receiving RGB-D data from the external camera firstly processed by the ROS wrapper for Openpose[5]. Then, the *posture analysis* module will receive the location of the body joints of the child's from the Openpose, and it could compute certain postural parameters related to handwriting, e.g., the trunk inclination angle. Lastly, posture-related information will be made available for the *iReCHeCk manager*.

### C. Dynamilis listener

The *Dynamilis listener* module provides information regarding the handwriting assessment and training activities performed in the *Dynamilis* app. Concretely, it relies on Firebase[6] APIs to connect to the Firebase database of the *Dynamilis* app and retrieve information of relevance every time the database is updated. Such information include, for the handwriting evaluation, the global score, the scores along the four sub-dimensions, as well as the values of the low-level features composing them. For the games, collected information include the type of game played by the child, the difficulty level and the score obtained in it, also expressed in the [0, 100] range. In both cases, the information is available for our system seconds after the completion of the activity.

### D. iReCHeCk manager

The *iReCHeCk manager* module, as the name suggests, is the information manager of the system, responsible for organizing the information about the child's status and performance provided by the various sensing modules in a homogeneous structure allowing for easy querying, analysis and export. At the same time, this module is responsible for keeping track of the phase the session is in and manage transitions from one to the other. Concretely, this is is implemented as a Finite State Machine (FSM), for which we rely on the smach ROS library[7], where states correspond to phases or sub-phases of the session and transitions are triggered by events detected from the sensing modules, or imposed by a human, as explained in Section III-F.

[2]All modules are freely available at https://github.com/irecheck/irecheck
[3]https://nuitrack.com/
[4]https://dynamilis.com/en/
[5]https://github.com/ravijo/ros_openpose
[6]https://firebase.google.com/
[7]http://wiki.ros.org/smach/Tutorials

### E. Decision Maker and User-Modeling

The *decision maker* module is responsible for the management of the session whenever the system is expected to run autonomously. The module relies on the run-time information made available by the *iReCHeCk manager* to suggest activities for the child. Currently, its mechanism relies on a hybrid system that combines simple conditional constructs to identify the handwriting dimension to train (via the related *Dynamilis* game) and short-term memory of the last three user's scores in order to break plateaus in the user's engagement.

The user modeling that drives the decision-taking process is implemented as another FSM, as represented in Figure 3, where we have the following states: *Single Win*, when the user wins after a previous non-winning play; *Single Loss*, when the user loses after a previous non-losing play; *Wining Streak*, when the user wins after a previous winning play; and *Losing Streak*, when the user loses after a previous losing play.

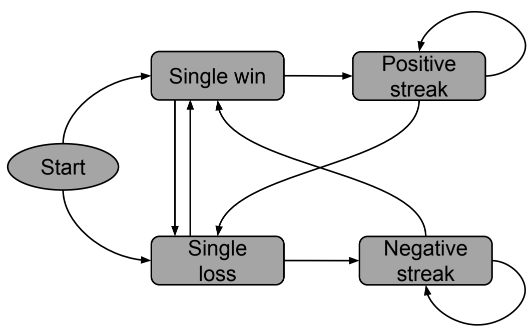


Fig. 3. Finite State Machine of our proposal for User-Modeling.

This modeling allows a short-term memory, in which by counting the times the participant is failing or succeeding at scoring in the last plays, the system triggers the type of the next game in order to avoid boredom for multiple failures or success. In this setup, what defines whether the player won or lost a play is if the score in the game is above (win) or below (loss) a predefined threshold.

The choice of the activity (game) is based on the corresponding lowest score in the last HWE. After 3 consecutive losses, it proposed a corresponding activity associated to the highest score in the last handwriting, which the child is expected to win. Similarly, after 3 consecutive wins, the robot suggested to play another activity, associated with the second lowest score in the HWE. The loss counter is not reset after changing the activity. So, if the user loses after a game change caused by a negative streak, it does not wait for 2 more failures to switch but does it immediately. On the other hand, the wining counter is reset after all game changes caused by a positive streak.

The awareness of such scenarios is important since small variations in recommendation systems might be as simple as effective in fighting boredom, as addressed in [21]. Furthermore, boredom has been shown to correlate negatively with learning in some learning activities, which often precedes disengaged behaviors such as off-task behavior [22].

For the future, we envision this module to allow different methods to base its decisions, for example, incorporating and building upon experts' knowledge in our decision system (Supervised Learning), and to implement personalization and adaptation over time (Reinforcement Learning algorithms) striving to balance maximizing the amount of time devoted to handwriting training and promoting a good posture, with the child's enjoyment of the interaction and wrapping up the session before fatigue and/or boredom can negatively affect it, as other than adapting the level of challenge in the games, decision-making systems were shown to be capable of further personalize and enhance robot-children interactions [23].

### F. Interaction Phases

Combining the presented resources of hardware and software, the *iReCHeCk* 's structure consists, on a high-level view, of four phases: (1) *Greeting*, (2) *Handwriting Evaluation (HWE)*, (3) *Training*, and (4) *Goodbye*.

*1) The greeting phase:* is triggered by the Nuitrack face detection algorithm and it is where the social robot welcomes the child. Expansions in this proposal to afford long-term memory in future works will allow the robot to bring up events and achievements of past sessions or to engage him/her in a short conversation, reinforcing personalization in the interactions.

*2) The Handwriting Evaluation phase:* is the part of the session where the robot invites the child to perform the Dynamilis handwriting assessment. The outcomes are used to guide the child in the games/exercises of the next training phase.

*3) The Training:* is a phase where the robot guides the child in the handwriting training through the games afforded by the Dynamilis app. The games - which serve as exercises - are chosen by the *Decision-Maker* node, based on the child's HWE outcomes and previous scores, as detailed in Subsection III-E.

*4) The goodbye phase:* is the last step of the interaction, where the robot wraps up the session, emphasizing on the achievements.

While the *greeting* and the *goodbye* phases are - intuitively - pinned in at the beginning and at the end of the session, the HWE and *training* phases can be used and combined in between them, as many times as one wants. Subsection IV-C shows an example of the combination of these phases, in a real use case.

## IV. EXPERIMENT

The analysis of the experiment is divided in two parts, as this preliminary test has a twofold goal: (1) assess the perception of the students about their experience with the system, and (2) assess the immediate impact of handwriting training activities on handwriting quality.

We consider the instantaneous analysis relevant for two reasons. First, prior to real-time assessments, it was impossible to be made since the handwriting quality assessment was performed by experts' observation, which implies interruptions or later assessment. Then, if there is a momentary peak of improvement in some of the mentioned handwriting dimensions, the robot could use this temporary variation in the interaction, e.g. to boost the child's self-esteem. Thus, we designed a series of handwriting evaluations alternated with personalized handwriting training.

### A. Participants

A total of 31 children (11 girls and 20 boys aged M = 8.52 years old, SD = 0.57) from two classes of grade three participated in the study[8]. The children come from diverse socioeconomic backgrounds and are all English speakers and writers. Only one had previously used the *Dynamilis* app. Two participants abandoned the experiment midway thus leaving us with the data of 29 participants for the analysis. Twelve sessions experienced network delays and in 8 of them we asked the children to restart the app. Children said that this request did not interfere in their interaction. In 4 cases we had to start the interaction manually instead of waiting the face recognition input. Nonetheless, participants also claimed it didn't interfere in their experience.

### B. Protocol

The sessions took place during the school time of the children and the setup was configured in an unused classroom, as shown in Figure 1. One at a time, participants were called out of their classrooms and briefed about the experiment by the researchers. The goal of this talk was to explain what participants should expect from their interaction with the robot; that they could ask questions or resign from the experiment at any time; and, finally, to explain the *Dynamilis* app functionalities and allow the participant to familiarize with it. The average time spent for this part was 10 minutes. After this first moment, the interaction with the robot started as described in Section IV-C. The average time for this part was 30 minutes. At the end of the session, the children were asked whether their had any doubt or question about what they just did and invited to answer a number of questions. This part lasted in average 5 minutes.

We asked 4 questions, the first 2 related to their feedback on their experience interacting with the system and the last 2 about their self-statement regarding enjoyment in participating in the activity. The questions for the participants were:

1) Did you understand everything the robot said?
2) What would you change in the experience you just had?
3) Would you like to participate again in this activity?
4) Did you enjoy participating in this activity with the robot?

Results are presented in Section V-A.

[8]This study has received ethical approval from the Human Research Ethics Committee of EPFL under protocol HREC 057-2021

### C. Interaction Setup

The interaction between the robot and the participant was designed combining the phases presented in Section III-F in the following sequence:

1) The greeting phase
2) A previous handwriting evaluation (HWE1)
3) The first training phase
4) An intermediate handwriting evaluation (HWE2)
5) The second training phase
6) A last handwriting evaluation (HWE3)
7) The goodbye phase at the end

The minimum time for the training phases was set to 5 minutes. This means that after 5 minutes of training, instead of suggesting the next game, the robot guided the student to the next phase, which in this case was always a HWE. The robot's social skills were programmed to be displayed through small movements in its arms and head, variation in the robot's facial display according to the sentence. For example, during the greeting phase it welcomed the child with a spoken utterance, a "Hello" gesture and a smiling face, and displayed a similar behavior in the goodbye phase. The robot was also reacting to the children's performance in the game by saying variations of "Well played", expressing a happy face and rising its arms when they succeeded, and by saying "Not a good result but I know you can do better than this", displaying a slightly sad face and lowering its arms when they failed. All the utterances of the robot included 3 possible variations on the same meaning, among which the robot randomly chooses in real-time, to prevent boredom as a result of repeating the exactly same line every time.

## V. Results

### A. Users Perception

Previous work have shown that questionnaires - such as Godspeed - often suffer from a ceiling effect caused by the novelty effect, especially when the users are children [15]. Thus, we decided for simple and straightforward questions to the kids in an interview mode, after their participation in the proposed activity.

For the first question, regarding their comprehension of the robot's instructions, 4 subjects (13,7%) have answered "no" for this questions. One of them said it was because of the noise (the very end of his session overlapped with one of the school's break), 2 of them said they have difficulties with the idiom (even though they have the regular classes in english) and 1 did not say why. All the others have answered "yes".

For the second question, about what they would change, 2 subjects (7%) said they would make changes in the robot (one said about the robot's speech - one of the children that claimed to not fully understand - and the other about increasing the robot movements), 4 subjects (13,7%) said they would skip the HWE, and only 1 subject suggested to make the games easier. All the other participants answered that they would not do changes in the proposed activity.

In the third question, regarding their desire to participate again, 5 subjects (17%) said that they were not sure whether they would do it again, 3 subjects (10%) gave affirmative answers with more words than just "yes" (First "I would love to" and the other 2 (7%) answered "for sure"). 2 (7%) said that they would perform again if they don't need to perform the HWEs. All the others said only "yes".

Finally, for the last question, regarding participants enjoyment, 2 subjects (7%) answered "no", one of them was the one who claimed to have issues with the idiom, 1 subject (3%) said he/she was "Not sure", 4 subjects (13,7%) gave affirmative answers with more words than just "yes" ("I enjoyed 20 out of 10", "I enjoyed 100 out of 10", "very much", "was super cool"). All the other participant said "yes".

The decisions of the system based on our simple user-modeling seemed to cover most of the scenarios. Only 2 (7%) participants felt on the worst case scenario of losing 2 times in a row, winning once and losing 2 times again. This is an undesired case because it keeps the student stuck in a situation they apparently can't handle, potentially leading them to frustration or boredom. This outcome may be supported by the fact that one of the students that felt in this case was one of the students that claimed to have not enjoyed the activity.

Moreover, two teachers from the two grades also experienced the activity with the system and gave us their feedback. In their opinion, the setup has a lot of potential to be used as consolidated tool for education, especially as a side activity in one of their rotating activity room, a special type of room with many educational booths where students are changing booths from time to time.

### B. *Study case of the Decision-Maker*

Figure 4 shows particular cases faced by the decision-maker system. The vertical dashes at activities 0, 7 and 15, represent the global scores the subject achieved in the HWE performed throughout the sessions while the colored lines are the progression of two subjects in the activities the system suggested according to the mechanism explained in Section III-E. Please, note that they had the same amount of activities but they could have different since the training sessions were time driven.

The red line represents the performance of a subject where the decision-making system faced issues trying to retake the user to a winning streak due to the users sequence of scores. In this case, the user got two failures in a row - it means performing a score equals or below 60 - (activities 1 and 2), and then succeeded once (activity 3), which resets the losing streak counter and prevented the system of changing the suggestion to another game, allowing the subject failing in the same game (activities 4 to 6) until a new HWE (activity 7) came where the student got the same suggestion. Then, after failing for 3 times (activities 8 to 10), the suggestion changed to a game in which the subject supposed to be good at (higher scores in the corresponding dimensions of HWE). After succeeded once (activity 11), the subject failed again 3

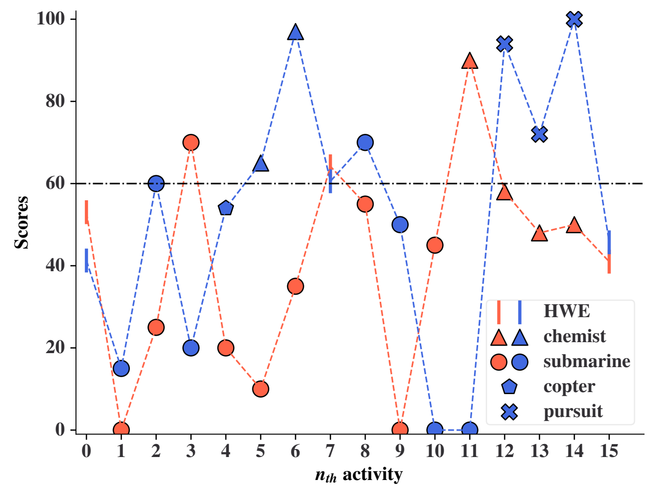


Fig. 4. The variation of Dynamilis activities in terms of scores and types for two children (colored by red and blue respectively) during the experiment. The shape of the markers represent different types of activity.

consecutive times (activities 12 to 14) but, instead of having a different suggestion, the time for the second training section went up and the system requested the subject to perform the final HWE (activity 15).

On the other hand, the blue line illustrates a case of success where, after the subject failed 3 times in the first suggested game (activities 1-3), the system suggested a game where the participant was supposedly better. The participant played the game "copter" instead of the suggested one that was "chemist" (activity 4). This misunderstanding shows the case in which participants may not understand the robot's instructions and follow to different ways than the expected one, which was out of the scope of this experiment. After being corrected by the researchers, the subject chose the correct suggested game, the subject got 2 successes (activities 5 and 6) until being interrupted by another HWE request (activity 7) where, again, the outcomes pointed out a necessity of increasing the pressure dimension of handwriting, which is addressed by playing the submarine. After succeeding once (activities 8), the subject failed 3 times (activity 9 to 11), and after getting a new suggestion, got higher scores (activities 12 to 14) until finish the second training session.

### C. *Immediate Variation Analysis*

The graph on Figure 5 represents the average scores of all the participants, along the dimensions used by Dynamilis to evaluate the handwriting quality, at the three HWE points in the activity.

By applying the two-sided Wilcoxon T test we observed no statistically significant difference between the averages of the scores from one HWE to the next, except for the *tilt* that showed a significant improvement between the first and the second HWE. Together with the fact that the children did not perform any handwriting-related activities before the first HWE phase, the increment in the tilt score might suggest that the proficiency of the handwriting skills in terms of pen

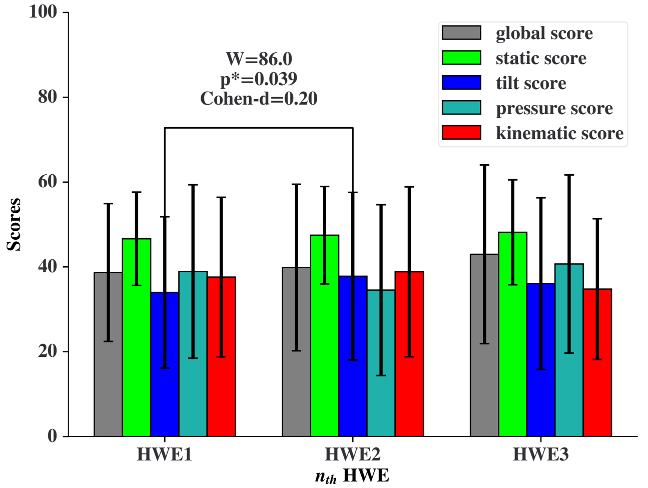


Fig. 5. Handwriting Evaluation (HWE) scores among three assessments.

tilt is volatile and the children need some time to recover to their normal proficiency.

These findings are valid for two main reasons. First, it confirms that one may not assume immediate improvement since even with temporary variations tend to stabilize afterward, reinforcing the fact that handwriting enhancement should target long-term practices. And second, small significant variations are possible, as shown by the tilt variation, and we can profit from it with the robot disclaiming them to the kids in envisioning sporadic boosts to their self-esteem.

## VI. Conclusions

In this paper, we presented the motivations and implementation of a system for handwriting training which uses a social robot as motivator agent. The proposed system architecture, discussed both in its hardware and software components, has been tested in a school environment. Albeit preliminary, the reported experiment proved that the proposed architecture enables the autonomous management of handwriting training sessions, mediated by a social robot. As a result, the system can be used to parallelize activities, as it allows autonomous training for students, enabling teachers to focus on other students.

## Acknowledgments

The authors would like to thanks the teachers Jessica, Fu and Frances that provided valuable support and feedback to the study.

## References

[1] B. C. Smits-Engelsman, A. S. Niemeijer, and G. P. van Galen, "Fine motor deficiencies in children diagnosed as dcd based on poor graphomotor ability," *Human movement science*, vol. 20, no. 1-2, pp. 161–182, 2001.

[2] R. I. Nicolson and A. J. Fawcett, "Dyslexia, dysgraphia, procedural learning and the cerebellum," *Cortex*, vol. 47, no. 1, pp. 117–127, 2011.

[3] T. Gargot, T. Asselborn, H. Pellerin, I. Zammouri, S. M. Anzalone, L. Casteran, W. Johal, P. Dillenbourg, D. Cohen, and C. Jolly, "Acquisition of handwriting in children with and without dysgraphia: A computational approach," *PLoS One*, vol. 15, no. 9, p. e0237575, 2020.

[4] C. Butler, R. Pimenta, J. Tommerdahl, C. T. Fuchs, and P. Caçola, "Using a handwriting app leads to improvement in manual dexterity in kindergarten children," *Research in Learning Technology*, vol. 27, 2019.

[5] D. Hood, S. Lemaignan, and P. Dillenbourg, "The cowriter project: Teaching a robot how to write," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, 2015, pp. 269–269.

[6] T. Gargot, T. Asselborn, I. Zammouri, J. Brunelle, W. Johal, P. Dillenbourg, D. Archambault, M. Chetouani, D. Cohen, and S. M. Anzalone, ""it is not the robot who learns, it is me." treating severe dysgraphia using child–robot interaction," *Frontiers in Psychiatry*, vol. 12, 2021.

[7] T. Asselborn, T. Gargot, Ł. Kidziński, W. Johal, D. Cohen, C. Jolly, and P. Dillenbourg, "Automated human-level diagnosis of dysgraphia using a consumer tablet," *NPJ digital medicine*, vol. 1, no. 1, pp. 1–9, 2018.

[8] R. Pintner, "A comparison of the ayres and thorndike handwriting scales." *Journal of educational psychology*, vol. 5, no. 9, p. 525, 1914.

[9] F. N. Freeman, "A new handwriting scale," *The Elementary School Journal*, vol. 59, no. 4, pp. 218–221, 1959.

[10] R. Bezzi, "A standardized manuscript scale for grades 1, 2, and 3," *The Journal of Educational Research*, vol. 55, no. 7, pp. 339–340, 1962.

[11] L. Hamstra-Bletz, J. DeBie, B. Den Brinker *et al.*, "Concise evaluation scale for children's handwriting," *Lisse: Swets*, vol. 1, pp. 623–662, 1987.

[12] W. Johal, "Research trends in social robots for learning," *Current Robotics Reports*, pp. 1–9, 2020.

[13] J. Kennedy, P. Baxter, E. Senft, and T. Belpaeme, "Social robot tutoring for child second language learning," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 231–238.

[14] J. E. Michaelis and B. Mutlu, "Reading socially: Transforming the in-home reading experience with a learning-companion robot," *Science Robotics*, vol. 3, no. 21, 2018.

[15] J. Nasir, U. Norman, B. Bruno, and P. Dillenbourg, "When positive perception of the robot has no effect on learning," in *2020 29th IEEE international conference on robot and human interactive communication (RO-MAN)*. IEEE, 2020, pp. 313–320.

[16] S. Frager and C. Stern, "Learning by teaching," *The Reading Teacher*, vol. 23, no. 5, pp. 403–417, 1970.

[17] S. Cooper and S. Lemaignan, "Towards using behaviour trees for long-term social robot behaviour," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '22. IEEE Press, 2022, p. 737–741.

[18] Z. Shi, T. R. Groechel, S. Jain, K. Chima, O. Rudovic, and M. J. Matarić, "Toward personalized affect-aware socially assistive robot tutors in long-term interventions for children with autism," *arXiv preprint arXiv:2101.10580*, 2021.

[19] K. Weber, H. Ritschel, I. Aslan, F. Lingenfelser, and E. André, "How to shape the humor of a robot-social behavior adaptation based on reinforcement learning," in *Proceedings of the 20th ACM international conference on multimodal interaction*, 2018, pp. 154–162.

[20] T. L. C. Asselborn, "Analysis and remediation of handwriting difficulties," EPFL, Tech. Rep., 2020.

[21] R. Warlop, A. Lazaric, and J. Mary, "Fighting boredom in recommender systems with linear reinforcement learning," in *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., vol. 31. Curran Associates, Inc., 2018.

[22] W. Hawkins, N. Heffernan, and R. S. Baker, "Which is more responsible for boredom in intelligent tutoring systems: Students (trait) or problems (state)?" in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 618–623.

[23] B. Schadenberg, M. Neerincx, F. Cnossen, and R. Looije, "Personalising game difficulty to keep children motivated to play with a social robot: A bayesian approach," *Cognitive Systems Research*, vol. 43, pp. 222–231, 2017.