CoWriter : Case Studies

Alexis Jacq[1,2], Séverin Lemaignan[1], Fernando Garcia[1], Pierre Dillenbourg[1], Ana Paiva[2]
[1]CHILI Lab, École Polytechnique Fédérale de Lausanne, Suisse,
[2]Instituto Superior Técnico, University of Lisbon, Portugal

ABSTRACT
Abstract comes here

Keywords
robot-supported educative activitiy, handwriting learning, learning by teaching

1. INTRODUCTION

2. THE COWRITER ACTIVITY

2.1 Children teach handwriting to the robot

2.2 Our approach

2.3 Learning and generating letters

2.4 robotic implementation

3. CASE 1 : DIEGO

3.1 Context

Diego is a five years old child. Her mother told us he had difficulties learning to write at school, particulary in drawing cursive letters. Before experiments, she provided us with a homework of Diego to show explicitly his handwriting level (fig).

From our perspective, Diego is shy and quiet. He suffers from a poor self-esteem much more than any actual trouble in writing.

3.2 Questions

The CoWriter activity needs a child engaged as interaction leader. In this study we consider the problem of long-term interactions: is it possible to sustain this engagement over several one-hour sessions?

3.3 Experimental settings

The experiment took place in our laboratory. Our goal was to provide Diego with an environment that would enable him to sustain engagement over four sessions of one hour, one session per week. We decided to introduce an appealing scenario that justified the activity to the child where a robot wants to learn handwriting. We used two Nao robots: a blue one (called Mimi) and an orange one (called Clem). Mimi was away for a scientific mission, and the two robots had to communicate by mails. But they decided to do it "like humans", with handwritten messages. While Mimi was good in handwriting, Clem had strong difficulties and needed the help of Diego.

The mission of Mimi consisted in the exploration of a mysterious hidden base. Each week, just before the session, it was sending a postal mail contening a picture, a curious object it found and a few handwritten words about its discoveries. The picture was representing itself exploring a dark room of the hidden base (that was actually our laboratory's workshop). The objects were 3D printed. In fact, there where puzzle pieces of a small 3D model of Nao robot but seen separately, it was not easy to guess it.

During the three first sessions, Clem (the other robot) was waiting for Diego with the received mail. It let Diego take a look at the picture and the object, and then it asked him to read the message. Finaly, Diego figured out a response and helped the robot to write it.

The fourth and last session was set as a test: Mimi, the "explorer" robot, had come back from its mission and it actually challenged Clem in front of Diego: *"I don't believe you wrote yourself these nice letters that I received! Prove it to me by writing something in front of me!"* This situation was meant to evidence the protégé effect: by judging the other robot's handwriting, Mimi would implicitly judge Diego's skills as teacher, and in turn, Diego's handwriting.

To complement the intrinsic motivation of helping a robot to communicate with another one, we gradually increased the complexity of Diego's task to keep it challenging and interesting (first week: demonstration of single letters; second week: short words; third week: a full message – Figure **??**).

Diego had to tell the robot what to write with small plastic letters (visible behind the robot on Figure **??**). A third person was here to send the formed word to the robot via the computer.

3.4 Results

Overall, Vincent provided 154 demonstrations to the robot, and he remained actively engaged over the four weeks. The story was well accepted by Vincent and he seriously engaged into the game. After the first week, he showed good confidence to play with the robot and he built affective bonds with the robot over the course of the study, as evidenced by some cries on the last session, and several letters sent by him to the robot *after* the end of the study (one of them 4 months later) to get news. This represents a promising initial result: we can effectively keep a child engaged with the robot for a relatively long period of time (about 5 hours).

No conclusion can be drawn in terms of actual handwriting remediation: we did not design this study to formally assess possible improvements.

However, as pictured on Figure **??**, Vincent was able to significantly improve the robot's skill, and he acknowledged that he had been able to help the robot: in that regard, Vincent convinced himself that he was "good enough" at writing to help someone else, and this is likely to have positively impacted his self-esteem.

4. CASE 2 : HENRY

4.1 Context

Henry, 5.5 years old child, is under the care of an occupational therapist. He has been diagnosed with visuo-constructive deficits. As an effect in writing activities, he was frequently performing random attempts and then was comparing with the provided template. What is more, Henry is restless and careless: he rarely pays attention to advice, even to what he is doing when he is currently drawing, and he is quickly shifting his attention from one activity to another.

Henry was working on number's allographs with his therapist. During a prior meeting, the therapist provided us with a sequence of numbers written by Henry **??**. Henry was sometime drawing horizontally-inverted allographs, mainly for "5".

4.2 Questions

This study focus on technical adaptations of the CoWriter activity for a child diagnosed with real writing deficits. Our objective is to investigate small modifications of the activity adapted to the troubles of Henry (visuo-constructive deficits and inattention) in order to maintain him focused on the activity during forty-minutes session, and to make the robot evidently learning from his demonstrations.

4.3 Experimental settings

The experiment was conducted in the therapist's surgery (four sessions spanning over 5 weeks). We assumed that a scenario like the one we used for Vincent was no longer relevant with Henry. We just introduced the robot and quickly said that it was seeking help to train for a robot handwriting contest.

In order to integrate our work with that of the therapist, we decided to adapt the CoWriter activity to teach numbers to the robot.

Since Henry was frequently drawing horizontally-inverted numbers, or even unrecognizable allographs, the learning algorithm of the robot was converging to meaningless scrawls. To fix this problem, we programmed the robot to refuse allographs that were too distant to a reference with a threshold we arbitrary fixed. In that way, the child was forced to take care on what he was providing to the robot as demonstration.

According to the therapist, it was easier for Henry to memorize the way to draw a number if it was always done is the same order, *e.g.* if the "5" was always drawn from the top-right tip down to bottom. Therefore we programmed the robot to refuse as well a good allograph drawn in a wrong order. But in order to reassure Henry about the right final allograph's shape, we made the robot able to recognize such a drawing, and, when it occured, to tell the child something like: *"Oh, this is exactly the shape of the number I want to learn, but can you show me how to draw it in the opposite order?"*

Also, to make the robot's progresses evident, we modified the initialization step of the learning algorithm to start with a roughly vertical stroke instead of a deformed number (round 0 on Figure **??**).

In this setup, we added a second tablet with one button per number. It was used by the child to chose a new number to teach to the robot. It also provided the possiblity to enter letters or words, and to switch to another activity (the robot telling a story).

4.4 Results

Despite his inattention, Henry was able to remain engaged in the activity during more than forty minutes in each session. In total, 55 allographs out of 82 provided by the child as demonstration were acceptable by the robot (with a progressive improvement from 13 out of 28 in the first session up to 26 out of 29 in the last session).

As soon as Thomas understood that the robot was only accepting well-formed allographs, he started to focus on it and he would typically draw 5 or 6 times the number before actually sending to the robot (the tablet let the children clear their drawing and try again before sending it to the robot). According to the therapist, it was the first time that Thomas would correct himself in such a way, explicitly having to reflect on how *another agent* (the robot) would interpret and understand his writing. Figure **??** shows how he gradually improved his demonstrations for some numbers, according to the metric we used to make the robot accept/refuse trials.

Since the robot's handwriting started from a simple primitive (a stroke), each time Thomas succeeded to have his demonstration accepted by it, the robot's improvement was clearly visible (as measured in Figure **??**). This led to a self-rewarding situation that effectively supported Thomas' engagement.

5. AUTOMATIC STUDIES

5.1 Context

The previous studies where adapted to children : we used a special design for each case in order to sustain the child engaged in the activity. This time, we conducted case studies with eight children trought a unique design. Those children have in common difficulties to learn cursive writing but the natures and intensities of those troubles are stricly different from one child to another. Valentine (7 years), Alexandre (6.5) and Jonathan (7) are under the care of an occupational therapist. Enzo (8) and Matenzo (7) are repeating their school year because of writing. Mona (6) and Adele (8) are bottom of their respective classes in writing activities. Nathan (7) is under the care of a neurologist, and has been diagnosed with specific language impairment. All of those children are expected, given their school level, to have in mind the allographs of cursive letters.

5.2 Questions

The main purpose of this study is to test the ergonomy of CoWriter. As we introduce the robot, we dont provide children with any scenario and give to them minimal explanations. Then we see how easily they take the role of the teacher and how seriously they try to help the robot.

5.3 Experimental settings

This experiment took place in the coffee room of a therapists shared surgery in Normandie, France. Over two weeks, each child came three time for one-hour session, except Adele and Mona who just did one session. A facilitator was here just to explain the rules of the game and tablets usage. As for Henry, the child had two tablets : one to choose a word (or a single letter) to teach, and one used by both the child and the robot to write. Sometime, if the child asked for, we provided him with allograph's template.

The starting point of the robot's writing was the same for all children: we used middle point between simple vertical strokes and letters. For this study, we wanted the robot to be only influanced by the demonstrations provided the child, so we did not projected allographs in an eigenspace. The generated trials of the robot where directly the middle way between demonstration and last state in cartesian space.

We added two buttons on the tablet interface. One green with a thumb up, and one red with a thumb down. Those buttons could be used by children to evaluate the robot (the green one was for rewards while the red one was for punishment). By this way, we could mesure the perception of the robot by the child: the more the child used evaluation buttons, the more he was playing the teacher, judging the robot instead of himself. It becomes possible to estimate if a child is playing seriously given correlation between his evaluation and robot's actual progression.

5.4 Results

All children maintained their engagement during the full sessions. They provided in average 42 demonstrations per session. All children used evaluation buttons and had preference to reward the robot (at the end, 99 rewards was accorded to the robot for 33 punishments).

Since sessions took place over only two weeks, we did not studied possible handwriting remediation in children.

We focused on correlation between children's evaluations and robot's progression. We estimated robot's progression as the difference between a starting score (score of the first robot's try when children has chosen a new word/letter to work on) and the current robot's score (after beeing taught by the child). that score is given by the average of euclidian distance between robot's try and a reference allograph over all letters of the word. Those references for letter allographs where beforehand drawn by us, taking inspiration in education.com cursive letters templates (http://www.education.com/slideshow/cursive-handwriting-z/). Let P_i be the estimated progression of the robot at time i. Of course, if the child chose to switch to a new word at time j , we got $P_j = 0$ and obviously $P_0 = 0$. To mesure how significant a child was rewarding the robot when it was progressing, we generated 10000 times the same number of reward/punishment but accorded at random times. Let R_i^n be the nth generated evalutation at time i ($R_i^n = 0$ if no evaluation occured at time i , $R_i^n = 1$ if a reward occured at time i and $R_i^n = -1$ if a punishment occured at time i), and $\overline{R}_i$ be the actual evaluation at time i. For each nth generated sequence of evauation, we compute a score of evaluation:

$$S^n = \sum_i R_i^n P_i$$

Then we can estimate the p-value p of the actual score:

$$\overline{S} = \sum_i \overline{R}_i P_i$$

given the distribution of the generated scores $(S^n)_n$, assumed to be gaussian:

$$p(\overline{S}) = \mathbb{P}\big[X > \overline{S}\big] = 1 - \phi\left(\frac{\overline{S} - \mu}{\sigma}\right)$$

where ϕ denotes the cumulative distribution function of standard normal distribution, μ the mean and σ the deviation of the generated scores $(S^n)_n$. As a result, we found that 5 of the 8 children obtained a score of evaluation significantly hight ($p(\overline{S}) < 0.05$). We reported score of evaluation p-values of each child in second-last column of Table 1.

Then, we also studied correlation between children's evaluations and their own progression. We did exactly the same analysis, using distances between children demonstrations and reference allographs to compute children progressions. Finaly, 3 of those 5 children that played "seriously" obtained score of evaluation of their own progression significantly hight (last column of Table 1). For those last children, it seems that the robot was reflecting their own performances, and while they were judging the robot positively (three times more rewards than punishments) they were actualy evaluating themselves.

child	demo	rew	pun	p(robot)	p(child)
valentine	127	24	6	2.4e-03	5.5e-02
enzo	223	20	9	1.7e-01	3.5e-01
matenzo	131	10	3	3.8e-03	7.9e-03
jonathan	98	10	5	1.5e-01	3.8e-01
nathan	115	16	4	5.3e-04	2.7e-03
alexandre	83	10	3	3.1e-02	6.0e-01
adele	35	4	2	5.0e-02	3.7e-02
mona	40	5	1	5.4e-01	2.0e-01

Table 1: results of evaluation. demo : number of demonstrations provided by the child over all session. rew : number of rewards accorded by the child. pun : number of punishments. p(robot) : how significant are the evaluation corresponding to robot's progression. p(child) : how significant are the evaluation corresponding to child's own progression.

6. DISCUSSION

7. CONCLUSIONS