Motivating Children to Learn Arithmetic with an Adaptive Robot Game

Joris B. Janssen[1], Chrissy C. van der Wal[1,2], Mark A. Neerincx[1,3], and Rosemarijn Looije[1]

[1] Netherlands Organization for Applied Scientific Research TNO, Kampweg 5, (3769 DE) Soesterberg

[2] Vrije Universiteit, Faculty of Sciences - dept. Computer Science, De Boelelaan 1081, (1081 HV) Amsterdam, The Netherlands

[3] Delft University of Technology, Faculty of Electrical Engineering, Mathematics, and Computer Science (EEMCS), P.O. Box 5031, (2600 GA) Delft, The Netherlands

Abstract. Based on a ‘learning by playing’ concept, a basic arithmetic learning task was extended with an engaging game to achieve long-term educational interaction for children. Personalization was added to this learning task, to further support the child’s motivation and success in learning. In an experiment, twenty children (aged 9-10) interacted three times, spread over days, with a robot using the combined imitation and arithmetic game to test this support. Two versions of the robot were implemented. In one implementation the complexity of the arithmetic progressed towards a predefined group target. In the other version the assignments increased in complexity until a personal end level was reached. A subsequent free-choice period showed that children’s motivation to play (and learn) was high, particularly when the game progressed to a personal target. Furthermore results show that most children in the last condition reach higher levels compared to the predefined group level.

1 Introduction

Children with a chronic lifestyle related disease have to take care of more aspects in daily life compared to their healthy peers. Support for these children in daily activities might therefore be beneficiary to them. The ALIZ-E project is aiming at a social robot for long-term interaction with these children. This robot should be able to perform three different roles over a relatively prolonged period of time: a *buddy* that provides a personalized and engaging interaction, an *educator* that teaches relevant knowledge and skills to ‘empower’ the child, and a *motivator* that persuades the child to adhere to a healthy lifestyle (e.g. the therapy, diet, medication) [10]. Several robot functions that support these roles are being developed incrementally, in an iterative process.

The overall scenario is based on a medical setting in Italy, where children diagnosed with diabetes will spend a complete week in the hospital after diagnosis to learn about the illness and the implications of it. For these (young) children one week away from home is a long time. The ALIZ-E robot intents to make

B. Mutlu et al. (Eds.): ICSR 2011, LNAI 7072, pp. 153–162, 2011.

the time in the hospital more pleasant, while supporting the education of the child's illness. Basic arithmetic skills help children with diabetes to count the carbohydrate intake for their nutrients. This ability can therefore contribute to young diabetics' self-efficacy.

In the project we use 'learning by playing' as a concept for the interaction. We combine the basic arithmetic learning task with an engaging game to achieve the project goal for long-term interaction. This last game is an imitation game, in which robot and child copy each others sequence of arm movements and, subsequently, add a new movement to this sequence. The robot attunes the number and repertoire of moves to child's performance based on principles of motivational feedback, in such a way that the children like to continue playing until they achieve their target [14]. This ensures that the child keeps being challenged which is an important factor in both intelligent tutoring systems [17] and game theory [6]. The imitation game balances the perceived challenges with the perceived skills of the child and proves to be challenging for the children. By additional personalization of the arithmetic assignments, we expect to further improve the child's motivation and learning performance. These effects are studied in an experiment.

How to measure motivation of young children is a non-trivial question. In addition to questionnaires and observations during the game, we will evaluate the motivation for interaction with the robot by providing a free-choice period [13] [16]. Furthermore, we choose to perform the experiment with healthy children, since we want to burden sick children as little as possible [5]. By ensuring that the general characteristics of the children in the experiment are similar to the diabetic target group (e.g. age), we expect to find principles that apply for both groups. In future experiments we plan to test this in a group of children with a chronic disease.

2 Aspects of Motivation

Literature distinguishes two types of motivation: intrinsic and extrinsic motivation. Our research objective is to establish long-term motivation, ultimately to make a change in behaviour possible. Extrinsic motivation, though effective for short-term task compliance, has been shown to be less effective than intrinsic motivation for long-term task compliance and behaviour change' [8]. We will therefore focus on *intrinsic* motivation. Fasola and Mataric [8] indicate several factors that contribute to intrinsic motivation, including praise, competition, real-time feedback of performance, optimal challenge, self-efficacy and self-determination. Vallerand et al. [16] describe several variables that decrease the intrinsic motivation and should therefore be avoided. These variables include: material rewards, surveillance, deadlines, lack of self-determination and negative performance feedback.

For this study, we used the imitation game. In this game, the robot and the child build sequences of arm movements together. Turn by turn the players repeat the existing sequence and add a new movement to the sequence. During the game, the robot gives motivational verbal feedback to maximize the performance of the child. Most motivational aspects were already incorporated into the imitation game: positive robot feedback (praise, real-time feedback on performance), no material reward for the child and the absence of deadlines or negative performance feedback. Other aspects were difficult to manipulate: competition, self-efficacy, self-determination. Optimal challenge is a aspect we have a closer look into: when a game is too easy, the player will become bored as opposed to the game being to difficult, which will result in the user becoming frustrated or anxious [8,6]. The optimal challenge is thus when there is a balance between perceived difficulty and perceived skills by the user. In the study presented here we implemented one version of arithmetic that should approximate the optimal challenge and one that does not.

3 Implementation

For this study we implemented and extended an imitation game with arithmetic assignments to mix fun and education. The new game is composed of two components: making arm movement sequences and solving arithmetic assignments. In line with Cohen et al. [2] emotional feedback in the game is attuned to match children's expectations. The game is presented to the children as a secret agent game, where arm movements are a secret code and to crack the code of the bad guys, the children have to solve arithmetic assignments. The performance on the two components of the game are not linked with each other.

Several worlds exist in the imitation game. The arm movements increase in difficulty depending on the current world (starting using one arm: 'left arm up' and extending towards both hands 'Both arms down'). Within a 'world', each player (the robot and the child) repeats the entire sequence and makes up a new movement, which is added to the sequence. The sequence ends when the length of the current world is reached or when the child makes a mistake. To prevent deception of the child, the robot does not make mistakes. The progress in 'worlds' is attuned to the performance of the child.

For the arithmetic implementation, 29 levels with 10 assignments are constructed. The levels have an increasing arithmetic difficulty (e.g. level 1: '$6 + 1$', level 10: '$41 - 10$', level 20: '4×42', level 29: '$1005 \div 67$'). The assignments are selected randomly within a level and displayed on a screen next to the robot (see Figure 1). The robot provides motivational verbal feedback after each answer.

Two versions of the robot game are implemented: one that has a predefined arithmetic group level as end goal and one in which children could reach the boundary of their arithmetic capabilities. Next to this distinction there is a difference in the learning algorithm between the two implementations. As long as no mistakes are made both versions have an increase of three levels each step, for fast convergence to an appropriate level. When a mistake is made in the group level version the level is increased by one from that moment on, resulting in a more moderate learning curve. For the personalized level implementation, a simple form of sensitivity analysis is used [15]. In the case of a mistake, the level is decreased by one, increasing the self-perception of arithmetic skills. Next to this, the levels are also increased by one from that moment on. Other motivational aspects including the arm movement part are not manipulated, in order to get a fair comparison between group and personalized level implementation.

The group goal is set at level 20, which was considered the appropriate level for children half way through year six (fourth grade in U.S.) based on information from Goffree [9] and Borghouts [1]. For the personalized level 29 is the maximum level, because the chances are slim that the children reach this level.

The robot contained a user model, which kept track of the movement progress and arithmetic level of each participant. The participant continues with the movements and assignments in the level they ended last interaction time.

4 Experimental Method

4.1 Participants

Participants were 20 Dutch children (11 F and 9 M, age 9 - 10 years) from elementary school 'Het Spoor' (Zeist, The Netherlands). This is a Jenaplan school, meaning that each child follows their own learning curve but still has to reach goals within a time frame. Balancing for their gender, the participants were randomly divided in two groups of 10 participants. All the parents/caregivers signed an informed consent.

4.2 Materials

NAO Robot. The robot used in this experiment was the NAO (Aldebaran Robotics, see Figure 2). The NAO was provided with a unisex name: Charlie. Charlie, and names with similar pronunciation, is an uncommon name in the Netherlands both for boys (494 in 2006) and girls (363 in 2006). We provided no clues about the gender of the robot, since we wanted to prevent the children being prejudiced to liking the robot because of its gender. Fluency TTS (v4.0, using neutral voice 'Diana') was used to generate the wav-files the NAO used. The software for executing the imitation game involves: a Wizard of Oz interface, a dialogue model, a user model, the arithmetic assignments database. The control software ran on various computers.

Wizard of Oz. In order to test the feasibility of components before complete implementation, we use a Wizard of Oz set-up. In this interface, the experimenter does the sensing (e.g. the wizard interprets the movements and speech of the children), initiates the dialog and controls the laptop that displays the assignments and the progress. The movements of the robot are also preprogrammed and the experimenter just has to press the right order of buttons to make the robot imitate the children. To the children it looks like the robot actually recognizes and remembers the set. At a later stage fully autonomous robot behaviours will be tested within the project.

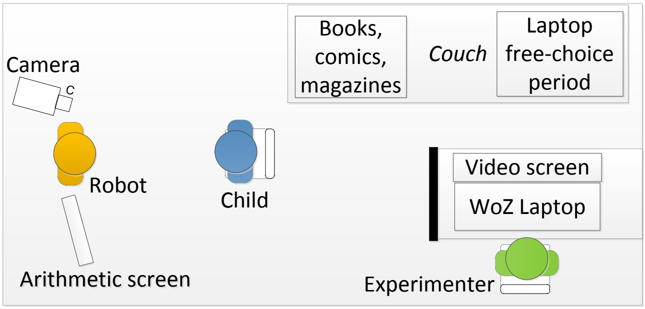


Fig. 1. The experimental setup

Fig. 2. The NAO robot

Experimental set-up. The experiment was conducted at the school in an office space. Unfortunately, there was no possibility for the experimenter to occupy a different room nearby, so the experimenter was in the same room as the child and the robot. The effects of the presence of the experimenter were minimized by placing a covering screen. The interaction was recorded on video. Figure 1 shows the experimental setup.

4.3 Experimental Design

The experiment performed had a between-subject design. The independent variable was the goal that could be reached in the arithmetic assignments, either the group or the personalized level. Between the two groups, one group interacted with a robot that adapted the level of the assignments to the child's performance and could proceed beyond the group level. The other group interacted with a group goal robot that followed a standard learning curve where the group level was the highest level that could be reached. The dependent variables were arithmetic performance and intrinsic motivation.

Measures. Two measures were used to measure the intrinsic motivation of the children to play the game with Charlie. A **questionnaire (subjective)** was constructed based on the Intrinsic Motivation Inventory (IMI)[1] (see other research [3][12][7]). Two of the seven subscales of the IMI were included: Interest/enjoyment (intrinsic motivation for playing game with the robot, 7 questions) and Relatedness (bond with the robot, 8 questions). The answers were measured using 7-point Likert scales (1 being negative and 7 positive). The original questionnaire in total and the separate subscales individually, were all validated. The questionnaire was translated into Dutch focused on children. The layout was altered for every session to keep children motivated to complete the questionnaire.

[1] http://www.psych.rochester.edu/SDT/measures/IMI_description.php

As an objective measure, the **free-choice period** [13] [16] was used. The free choice period was a period of five minutes in which the child could choose what to do: keep playing with the robot, read children's comics or do interactive Internet learning games on a computer. The time spend interacting with the robot was measured and functioned as an objective measure for the intrinsic motivation of the child to interact with the robot.

4.4 Procedure

During a short introduction in class, the children were able to see the robot beforehand. For the individual interaction moments, the experimenter introduced the child to the robot and explained the course of the experiment. Each interaction session lasted about 20 minutes, based on the average attention span for children of this age [4]. The child played the game with the robot for about 15 minutes. The game was ended after the 13th minute at a natural moment when the level was completed, resulting in a 13 to 17 minute interaction time. Afterwards the free-choice period was started by the experimenter. The researcher stated that the experiment had ended and that the child had 5 minutes to do as it pleased, choosing between the mentioned options as long as it stayed inside the room (options were presented in a random order). Finally, the child completed the questionnaire. The experiment was performed three times for each child over the course of two weeks. The rationale behind this was to experiment with the constructed user model and to overcome the initial enthusiastic response displayed by children when first meeting the robot. To reward the children, they received a picture of themselves with the robot. The school received technical Lego and was given a robotics lesson for the class after all sessions were completed.

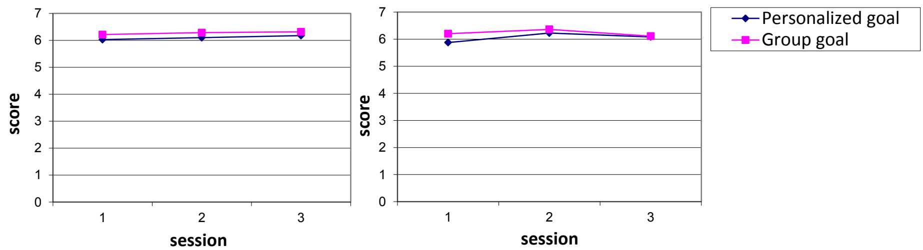
5 Results

5.1 Motivation

First the quantitative results will be discussed, based in the two different motivation measures used in the experiment.

Questionnaire. The results of the intrinsic motivation questionnaire are analyzed for each session. The answers represent the motivation to play the game with the robot and the bond with the robot. From the analysis the participant that did not start the third session is excluded and missing data is filled with a random participant from this condition to make ANOVAs possible.

We expected that the children in the personalized goal robot condition would score higher on the motivation scale than the children in the group goal condition. Results show that both scales are rated high (see Figure 3). The standard deviations are small, they ranged for Interest/Enjoyment from 0.39 to 0.70 and for Relatedness from 0.37 to 0.66. For Interest/Enjoyment the repeated measures ANOVA over the runs has as result: $F(2, 48) = 0.01$, $p=0.99$. The result



(a) Interest/Enjoyment (b) Relatedness

Fig. 3. Questionnaire results

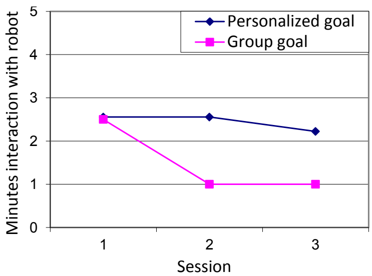


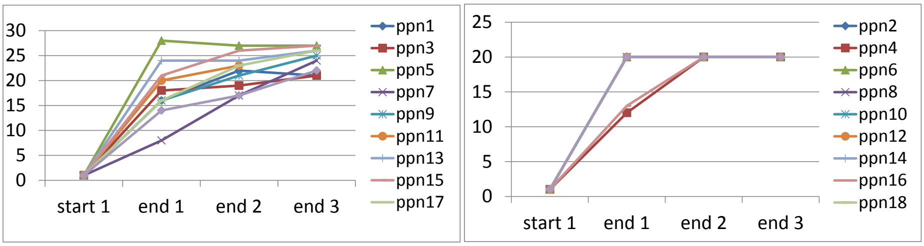
Fig. 4. Amount of time spend with robot during free-choice period

Table 1. Results of the free-choice period. Time interacting with the robot (in mm:ss). * entails the child stopped the interaction before the free choice period started, ** entails the child was absent

Personalized goal robot				**Group goal robot**			
Child	**Run1**	**Run2**	**Run3**	**Child**	**Run1**	**Run2**	**Run3**
1	*	*	0:00	2	0:00	0:00	0:00
3	2:20	3:13	5:00	4	5:00	5:00	5:00
5	5:00	0:00	0:00	6	5:00	0:00	0:00
7	0:00	5:00	5:00	8	0:00	0:00	0:00
9	0:00	0:00	0:00	10	5:00	0:00	0:00
11	5:00	5:00	**	12	5:00	0:00	0:00
13	5:00	5:00	5:00	14	0:00	0:00	0:00
15	0:00	0:00	0:00	16	5:00	5:00	5:00
17	5:00	5:00	4:30	18	0:00	0:00	0:00
19	0:00	0:00	0:00	20	0:00	0:00	0:00
AVG	2.6	2.6	2.5	AVG	2.5	1.0	1.0

for the ANOVA for the Relatedness questionnaire is $F(2, 48) = 0.16$, $p=0.85$. Thus both questionnaires do not provide significant differences between the two conditions.

Free-choice period. Table 1 shows the amount of time spend with the robot per participant for each session during the free-choice period and Figure 4 shows the means graphically. In the free-choice period following the first interaction the time spend with the robot is about the same (mean personalized = 2.6min, mean goal = 2.5min). This was expected beforehand, due to the new experience of the interaction with the robot. Video footage shows that the children were in general very excited to play with the robot. After the first session, the results started to differentiate between the two conditions. Most children that interacted with the personalized goal robot continued to play with the robot during the free-choice period, whereas the children that interacted with the group goal robot displayed,



(a) Personalized goal **(b)** Group goal

Fig. 5. Performance for arithmetic assignments

on average, a decline in the amount of time spend with the robot during the free-choice period. Child 1 stopped the experiment before the free-choice period and child 11 was absent during the third session resulting in missing data points.

Because the results are not evenly distributed, we ran a nonparametric Mann-Whitney U-test to establish whether the differences between the two conditions are significant. A one-tailed Mann-Whitney U-test showed that the difference was significant ($p < 0.05$) in favour of the personalized goal robot.

5.2 Results Arithmetic Aspects

We expected the children to reach arithmetic level 20, which corresponded with half way through 6th grade. However figures 5a and 5b show that the children that interacted with the personalized goal robot, performed above the expected norm on the arithmetic assignments (average 24.7). Especially child 5 stands out in arithmetic skills. The graphs show that most children already reached level 20 after the second interaction. From these results, we can derive that most children participating in the experiment are ahead in their arithmetic education and that playing with a personalized goal robot makes sense, since the individual levels differ from each other.

We looked into the interaction between the free-choice period and the performance on the arithmetic assignments. When Figure 5a and Figure 5b are linked with Table 1, it shows that the two children who played with the group goal robot during the free-choice period after session 2 and 3, were actually the children that did not reach level 20 after the first interaction. It appears as though the continuing increase in level motivated the children to play with the robot during the free-choice period. When looking at the children that played with the personalized goal robot, we see a similar trend. Child 5 performed very well on the assignments and reached his personal level during the first session. During the free-choice period the child chose to read instead of playing with the robot. However, Child 13 who also reached his personal level at the first session, did continue playing with the robot during the free-choice period. Hence, some children who perform at top level still like to play with the robot.

6 Conclusion and Discussion

In this paper we present a study that builds upon the principles of learning by playing. By combining a basic arithmetic task with an engaging game, we create a robot game for children. In an experiment we look whether personalization of the learning task has an effect on children's motivation and learning. In general we found that the children are very motivated to play the game with the robot. The motivation stays at a high level for all three interaction moments. The objective motivation, free-choice period, stays high when they interact with a robot offering a personalized learning goal. Most children who play with the personalized goal robot keep interacting with the robot the full five minutes of the free-choice period and the two children who are a bit slower to reach level 20 in the group goal session keep interacting with the robot during the free-choice period.

The personalized goal version shows that the group goal is not high enough for most of the children to reach their maximum capabilities. The group goal is thus not challenging. In sum, this robot game provides a promising approach to support long-term interaction even when the interaction is not all about fun. This is promising for the use of a social robot for long-term interaction with diabetic children. In a next study, diabetic children will participate in the study to see if the results can be reproduced with this specific population.

From a methodological perspective, the free-choice period proves to be very useful to study motivation effects with children. It appears that children answer the questions socially desirable. Despite several urges of the experimenter to rate how they really feel about the game, children seem to stay away from the 'negative' answers even though some children seem sometimes a little bored during the game (based on video footage). In future, we plan to include more detailed observations on communication behaviour, like eye-contact (gaze wondering off). In addition, we will improve the questionnaires. For example, research on Likert scales for children suggests to use a 3-point scale [11].

Acknowledgments. This work is (partially) funded by the European Union FP7 ALIZ-E project (grant number 248116). Furthermore the authors would like to thank the teachers and the children of 'Het Spoor' (the school) for their participation in this study.

References

1. Borghouts, C., Buter, A., Dekker, J., Hoogenberg, E., Kopmels, D., van Oostenbrugge, M.: Interactie in Rekenen. In: Bazalt Educatieve Uitgaven (2005)
2. Cohen, I., Looije, R., Neerincx, M.: Children's Recognition of Emotions in Robots' Face and Body. In: Proceedings of the 6th International Conference on Human Robot Interaction (HRI 2011), Lausanne, Switzerland, March 6-9 (2011)
3. Colombo, R., Pisano, F., Mazzone, A., Delconte, C., Micera, S., Carrozza, M.C., Dario, P., Minuco, G.: Design strategies to improve patient motivation during robot-aided rehabilitation. Journal of NeuroEngineering and Rehabilitation 4(3) (2007)

4. Cornish, D., Dukette, D.: The Essential 20: Twenty Components of an Excellent Health Care Team. Rosedog Pr (2009)
5. Coyle, D., Doherty, G., Matthews, M., Sharry, J.: Computers in talk-based mental health interventions. Interacting with Computers 19(4), 545–562 (2007)
6. Csikszentmihalyi, M., Abuhamdeh, S., Nakamura, J.: Handbook of Competence and Motivation. In: Flow, ch. 32, pp. 598–698. The Guilford Press, New York (2005)
7. Deci, E.L., Eghrari, H., Patrick, B.C., Leone, D.: Facilitating internalization: The Self-determination theory perspective. Journal of Personality 62, 119–142 (1994)
8. Fasola, J., Matarić, M.J.: Robot Motivator: Increasing User Enjoyment and Performance on a Physical/Cognitive Task. In: Proc. IEEE International Conference on Development and Learning (ICDL), Ann Arbor, MI, USA (2010)
9. Goffree, F., Oonk, W.: Reken Vaardig: op weg naar basale en professionele gecijferdheid. In: Noofdhoff Uitgevers B.V. (2004)
10. Looije, R., Neerincx, M., de Lange, V.: Children's responses and opinion on three bots that motivate, educate and play. Journal of Physical Agents 2, 13–20 (2008)
11. Markopoulos, P., Read, J., MacFarlane, S.: Evaluating children's interactive products: principles and practices for interaction designers. Morgan Kaufmann Pub. (2008)
12. Plant, R., Ryan, R.M.: Intrinsic motivation and the effect of self-consciousness, self-awareness, and ego-involvement: An investigation of internally-controlling styles. Journal of Personality 53, 435–449 (1985)
13. Reeve, J., Deci, E.L.: Elements of the Competitive Situation That Affect Intrinsic Motivation. Society for Personality and Social Psychology 22(1), 24–33 (1996)
14. Robben, S.M.B.: Graduation Thesis: It's NAO or Never! Facilitate Bonding Between a Child and a Social Robot: Exploring the Possibility of a Robot Adaptive to Personality (2011)
15. Russell, S., Norvig, P.: Artificial Intelligence, a modern approach. Pearson Education, New Jersey (2003)
16. Vallerand, R., Gauvin, L., Halliwell, W.: Negative Effects of Competition on Children's Intrinsic Motivation. The Journal of Social Psychology 126(5), 649–657 (1986)
17. Woolf, B.: Building intelligent interactive tutors, vol. 3. Elsevier (2007)