Universiteit Leiden

Master Computer Science

The effect of robot role and reading comprehension on children's learning outcomes

Name: Xueyu Zhao
Student ID: s2154293

Date: 24/01/2021

Specialisation: Science Communication and society

1st supervisor: Joost Broekens
2nd supervisor: Rifca Rijgersberg
2nd reader: Peter van der Putten

Master's Thesis in Computer Science

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Abstract

In this study, we investigate what is the effect of robot role is on learning outcomes, and whether reading comprehension is mediating this potential effect. The pedagogical robot takes on either a tutor role who teaches and tests the children, or a learner role that learns from the children to reinforce what he/she knows. Learning outcomes are divided into two parts: cognitive learning outcomes and affective learning outcomes. We not only paid attention to children's learning gains, but also to their emotions. To investigate how the two different roles impact children's learning outcomes, we designed a between subject experiment. Thirty Chinese Dutch Children aged 7-13 years old from a Chinese school participated in a chemistry learning activity with one of the two roles of robot. The results from our study revealed that roles have no significant effect on children's cognitive and affective learning outcomes. However, we found that children with different level of reading comprehension ability performed differently in the posttest. After analysis, it turns out that reading comprehension ability does have an effect on children's cognitive learning outcomes under the influence of the reading materials provided in this experiment. Children with high reading comprehension ability performed better in the posttest than those with low reading comprehension ability, which means that they learned more in this activity. This research of children-robot interaction can provide useful insights for educational robot role designs.

Acknowledgements

This master thesis was done at LIACS with my supervisors Joost Broekens and Rifca Rijgersberg. I would like to thank them for their help, advice and supervision during this project. Also, a lot of thanks to Peter van der Putten for being my second reader. What's more, I have done this research during a very tough corona period and ran into a lot of difficulties recruiting participants and doing field experiments. I want to thank Matthijs van Leeuwen' s help with my study plan, and the Chinese school's principal Wan for the help of organizing the experiment. And I want to thank Igor for helping me translate all English materials to Dutch.

Table of Contents

1. Introduction

Robots have potential to be used as a new technology for education [8]. Educational robots are designed to play different roles such as tutors [2], learners [3] or motivators [4] to deliver learning experiences through social interactions with their users [1]. Those robots perform well-defined tasks with pre-programed behaviors according to those roles. Many researchers have studied the robots' roles [5, 6, 7]. From the perspective of role design or the perception of children on those pre-designed robot roles, research has been done to study the interaction between robot and users, and also the user's learning outcomes [1]. Nevertheless, not only the robot designs are worth paying attention to, but also the users' characteristics are of importance. In this research we investigate the effect of role and the interaction between role and reading comprehension as a user characteristic.

1.1 Robot roles

Roles are being used in experimental human-robot interaction [9, 10]. Robots can play different roles according to different tasks and contexts. According to our survey of instructional style [see appendix 5], teachers adjust their instructional behaviors to meet their children's learning needs. And if we want to design good educational robots, they need to be designed based on existing human educational research. In other words, robot behaviors are usually mimicking those human teacher instructional styles. In this study, we picked two robot roles (tutor and learner) to interact with our participants. They are well defined by the previous studies of pedagogy and robotics. To match these two roles with suitable children who have specific characteristics, we based our research on the Aptitude-treatment instruction theory, which showed that a learner's abilities is of importance when choosing the ideal instructional approach [22]. It states that high structured instructional teaching turns out to be most successful with children of lower ability; on the contrary, low structured teaching results in better learning for high ability learner [22]. Therefore, a competent tutor robot with structured instructions and an incompetent learner robot with low structured instructions are designed for this study. In this section, I provide an overview of the two different roles and how the robot should behave according to its role. This research studies the effect of these two different robot roles on children's cognitive and affective learning outcomes.

Tutor robot
Robots can be a *Tutor* for humans. According to Belpaeme et al., children tend to gain new knowledge and skills from knowledgeable partners [16]. Therefore, tutor robots are designed to show a high level of competence, which means that the robot knows a lot. It would teach and test the children. This type of robot is often designed for young children's curricular domain [1]. Early field studies placed robots into classrooms to observe whether they would have any qualitative effect on the learners' emotions and learning process, but current research tends to be conducted in laboratory settings and classrooms, which is easier to control the variables [11].

Learner robot
Belpeame's study shows that children are able to gain a lot of educational benefits from a learner robot [38]. This type of robot acts as if they are low in competence, which means they are

unknowledgeable and make mistakes very often. It asks silly questions and look for help from the children. Then the children try to explore answers to the robot's questions. During the exploration, children could reinforce what he/she knows. This is an instance of learning by teaching, which is widely known in human education, also referred as the protégé effect [13]. This process involves the children trying to teach the robot. Studies placed robots into classrooms [19, 21] and laboratory settings.

1.2 Learning outcomes

This study will focus on the cognitive learning outcomes and affective learning outcomes of the children who interact with the robot. Discussion during the 1984 Convention of the American Psychological Association led Bloom and a group of educators to classify educational goals and objectives [12]. Eventually, it resulted in a taxonomy of three domains: cognitive, affect and psychomotor dimensions [12]. In this study, we focused on the cognitive and affective aspects. Firstly, according to Bloom's Taxonomy model [12], thinking could be classified according to six cognitive levels of complexity: remembering, understanding, applying, analyzing, evaluating and creating [12] [(see Fig.1). In our experiment, we studied the second level of cognitive learning outcomes which focus on how children learn from oral, written, and graphic information through interpreting, exemplifying, classifying, summarizing, inferring, comparing, and explaining [33]. Learning gains were measured by applying pre- and posttests of children's knowledge.

Secondly, affective outcomes refer to qualities that are not cognitive learning outcomes [39], for example, the learners' emotions, their feelings when they interacting with the robot. Affective outcomes are more varied and can include self-reported measures and observations by the experimenter [39]. In this study, the affective learning outcomes were measured by the self-Assessment Manikin, which is an efficient cross-cultural measurement of emotional response [34]. This questionnaire uses the pleasure, arousal, and dominance (PAD) model to measure people's emotions during this activity [34].

Besides the cognitive and affective learning outcomes, how the participants perceive the level of competence and whether they actually think that the robot is a tutor or a learner were also measured by asking them adjective-based questions designed by Peters et al., in our questionnaire [3].

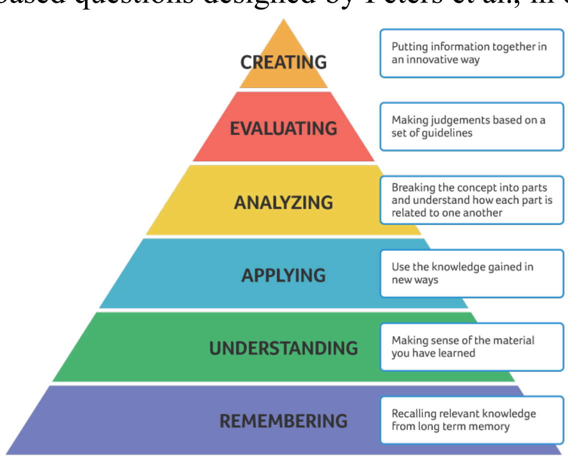

Fig. 1. Bloom's Taxonomy.

1.3 Adaptive robots

Previous studies have shown that adaptive robots could lead to better learning outcomes for the following reasons: Firstly, robot that adjusted its level of challenge provided to the children could motivate them to learn [8]. According to Janssen et al., the level of the assignment was adjusted according to the performance on the assignment of the child. As a result, the child had more motivation to learn when they were assigned to an adaptive robot [8]. Secondly, the robot that adapted itself to provide help was shown to promote the children's social skills [9]. In this study, the robot varied its behavior (adjusted gesture, gaze, head shift and voice) according to the characteristics of each child based on the cues from the child's head movement. As a result, the children who had autism's social skills were developed [9]. For the reasons above, an adaptive robot has the potential to help children learn better. Therefore, it is of importance to know the influence of different factors on the learning outcomes because it allows us to better adapt robots' behaviors according to the child in the future.

Como and Snow distinguished two level of adaptive teaching: Macro and Micro [27]. Macro adaptions are general instructional adjustments that are made based on information such as formal assessments [27]. Mirco adaptions are adaptions made within the moment-by-moment outcomes [27]. In this study, we design a macro level adaptation for a teaching robot. We collected information of children's reading comprehension ability from their teachers. Further, human teachers' teaching styles are applied to a robot based on pedagogical theories. We studied the effects on learning outcome of robot role, as well as the interaction effects with reading comprehension as a characteristic of the child.

2. Research question

Main question: What is the effect of robot competence expressed as a tutor versus a learner role on children's cognitive and affective outcomes, and what is the interaction effect with reading comprehension?

Sub-questions:
(1) What is the effect of robot role on children's cognitive learning outcomes?
(2) What is the effect of robot role on children's affect learning outcome?
(3) What is the effect of role adaptation when it comes to cognitive and affective learning outcomes compared?
(4) What is the effect of reading comprehension ability on impact of roles on the cognitive learning outcomes?

Hypotheses

(1) Robot role has a significant effect on children's cognitive learning outcomes.
(2) Robot role has a significant effect on children's affective learning outcomes.
(3) Role adaption has a significant effect on children's cognitive and affective learning outcomes.
(4) A learner role results in higher learning outcomes compared to a tutor role for children with high reading comprehension ability.

(5) A tutor role results in higher learning outcomes compared to a learner role for children with little reading comprehension ability.

3. Methodology

3.1 Protocol

To investigate our hypotheses, we conducted a between subject child-robot interaction experiment. Each child interacted with one of the two roles of the robot. Before the interaction, children provided some basic information and did a pre-chemistry test. During the interaction, children learned with a robot about atoms, molecules, periodic table and daily chemistry. Children's behaviors and emotions were observed and documented by the experimenter. Robot role was varied between subject. After the interaction, they were given another questionnaire, which included a knowledge post-test and some additional questions that aim to understand children's affective learning outcomes and their perception towards the robot.

Before doing the main experiment, we conducted a pilot study. Both experiments are reported upon in section 4.

3.2 Materials

Robot NAO
Nao robot, a 54-cm-tall humanoid robot developed by Softbank Robotics has 25 degree of freedom, and is widely used in educational robot studies (see Fig.2). Nao's humanoid appearance is suitable to play either a tutor or a learner role. It can move, gesture, talk and perform basic perception and behaviors.

Fig. 2. Nao robot, Picture from SoftBank Robotics.

Robot role design

Tutor robot and learner robot manifested their role with verbal and non-verbal behaviors. The tutor robot created a very structured learning environment by giving presentations and explanations to the children. For the verbal behavior, we have written a script for tutor robot that showed its high competence [see appendix 1]. This script mainly had four parts talking about atoms, molecules, modern periodic table and daily chemistry. In each part, the tutor robot provided the participants with knowledge and information first, and then it would ask a question accordingly. When the children answered correctly, the robot gives affirmation. When the children give the wrong answer, the robot would correct them and give encouragement to the children. For the non-verbal behavior, according to Peters et al., robots that have a stable body posture, fixed gaze at audiences and frequent hand gestures are perceived as high in competence [32]. Therefore, we designed our tutor robot accordingly (see Fig.3).

For the learner robot, it acted like an unskilled partner for the target children to teach. This teachable agent intentionally made mistakes. Through this capability, the robot could be taught by the children, who themselves may learn through their teaching [29]. For the verbal behavior, we also wrote a script that about the same chemistry content, but the robot acted like it knew nothing by constantly asking the children chemistry questions. For instance, the robot would ask, "Oxygen is the lightest element in the periodic table. Am I right?" The children needed to teach the robot by answering its questions. We wanted the children to learn during this teaching process. For the non-verbal behavior, we designed this robot to have low frequent hand gestures [32]. In this study, the robot had two versions of script, English and Dutch. The robot was programmed using the robotsindeklas.nl platform [28].

Behavior / Role	Verbal	Non-verbal
Tutor	Explaination Presentation Ask questions and give feedback	Stable body posture Fixed gaze Frequent hand gestures
Learner	Asking for help Making deliberate mistakes	Non-stable/wobbling body posture Low frequent hand gestures

Fig. 3. Verbal and non-verbal behaviors of tutor/learner robots.

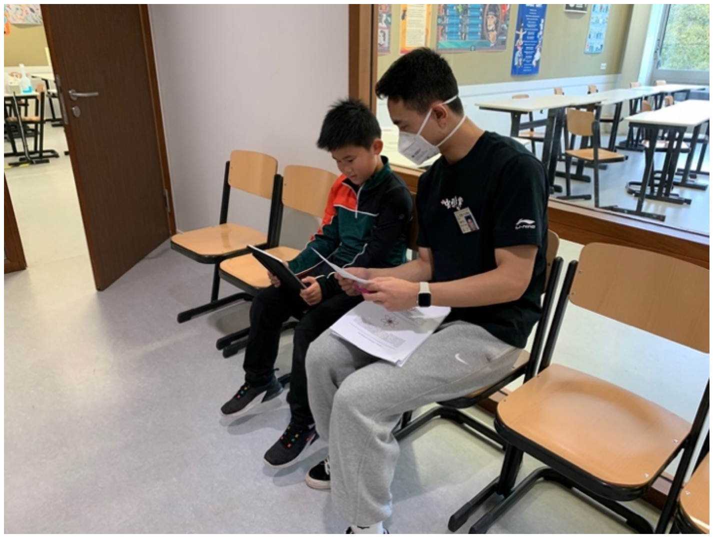
Fig. 4. Participants read the material with iPads.

Reading material

Reading material about atoms, elements, molecules and daily chemistry was prepared for children to read before the interaction [see appendix2]. In the first paragraph, children got the information that atoms are the smallest unit of an element that still maintain the chemical properties of the element. In the second paragraph, children should know that molecules are made up of more than one atom. In the third paragraph, children were given a modern periodic table. In the last paragraph, interesting chemistry facts were demonstrated. Children were asked to read the three-page material after finishing the pre-test and before entering the experiment room (see Fig.4). It's for them to gain some knowledge about this activity so that they can answer the questions of either the tutor robot or the learner robot. During the interaction with the robot, children were allowed to look for information in the reading materials.

Two questionnaires

The pre-and post-tests were aimed at testing their knowledge about atoms, elements, molecules etc. The pre-questionnaire aims to get the basic information of the participant (name, age, gender, experience with robots and learning habits), The post-questionnaire will evaluate children's perception of the robot and their emotions during the interaction. [see appendix3].

4. Experiments

We conducted three experiments, including a pilot study, a main experiment and an additional experiment. In the pilot study, we wanted to test our experiment procedure and get feedback from the participants. The main experiment was conducted in a Chinese language school where a lot of children learn Chinese on the weekend. We collected the data of children's interaction with the robot and observed their behaviors and actions. During this main experiment, we found that the content of this activity was very different for children between 7-13 years old. Therefore, we arranged an additional experiment. This experiment's setting remained the same, but our participants were adults.

4.1 Pilot study

We conducted a pilot study in order to see the robot's performance and eliminate potential problems through the whole experiment.

Participants
Three Dutch children were invited to Leiden University's robot lab, two boys (nine-year-old and eleven-year-old) and a girl (eleven-year-old). Those children have had experience interacting with NAO robots, so they were neither too supervised nor too scared when seeing the robot.

Procedures
Three participants were seated at a table. In front of them were cards, reading materials, a computer showing slides and a Nao robot. They were asked to read the materials before the interaction. After they finish reading, they interacted first with a tutor robot and then a learner robot. Both robots talk about four parts including atoms, molecules, the modern periodic table and daily chemistry.

During the pilot study, we found things that could be improved. Firstly, children were overwhelmed by all the materials in front of them. They didn't know which one to start with and how to move to the next step. Therefore, we decided to present one material at a time in the actual experiment to keep things in order, and to help children better understand our instructions. Secondly, there are some language errors in the robot's speaking, which was also fixed immediately after this pilot study. Furthermore, we interviewed those three children and based on their feedback we decided that some parts of the reading materials were a little bit long and too hard for them to understand. So, we decided to refine the reading materials as well.

Improvement
- One material offered to the participant at a time during the experiment
- Language errors fixed
- Refined the reading material

4.2 Main experiment

Participants
Thirty children aged between seven and thirteen (Male: 16, Female: 14) participated in this experiment. Most of them have no experience playing with a NAO robot (27 children have no experience, one child has experience and two children have experience with other robots but not NAO). As for their learning habit, eleven children would like to learn things on their own, and nineteen children self-described as preferring to learn from their teachers first. Among these children, ten children were evaluated to have good reading comprehension ability by their teachers, and twenty children were evaluated less good at reading. It's worth mentioning that we conducted this experiment at a Chinese school. Most of the children are Chinese-Dutch (they are fluent in speaking Dutch, but they tend to learn Chinese in the weekend). Besides the language preference, their culture background is worth noticing. They have a mixed mindset of both Chinese and Dutch culture, which could influence their behavior and thinking.

Design

We followed a between subjects design investigating the effect of robot role (tutor, learner) on learning outcomes (cognitive outcomes and affective outcomes) of children with reading comprehension as covariate.

Procedures

Firstly, an experimenter asked the participants to wait outside of the room and fill in questionnaire Part 1. Then, they needed to read the materials. Next, children went to another classroom one by one in numbered order. Then, the experimenter briefly introduced Nao robot to the participants and started the interaction. While they were interacting with the robot, the experimenter observed and took notes of children's behaviors and actions. After the children finished the interaction, they left the experiment room and sit in a chair to finish questionnaire Part 2. Once they finished it, they were given a gift for their participation (see Fig.5,6).

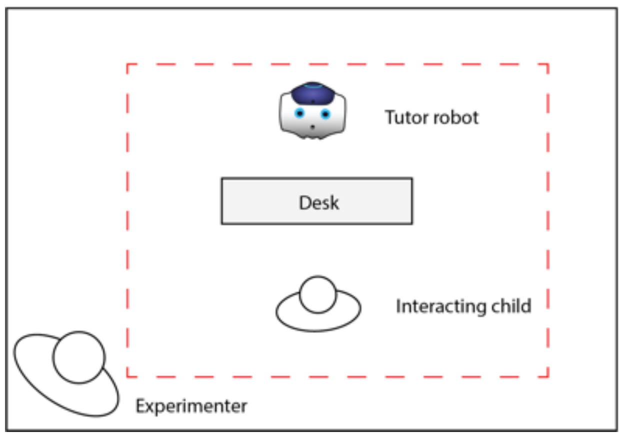

Fig. 5. Physical setup of the system within the classroom.

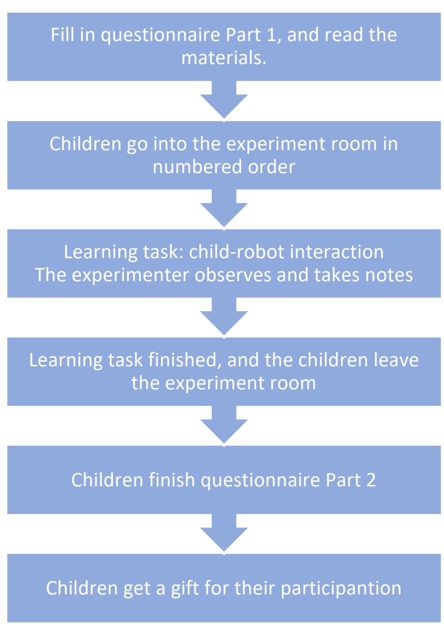

Fig. 6. Experiment process.

Measures

Reading comprehension ability

Children's reading comprehension ability was measured by their teachers. They evaluated the children with two levels: high or low, according to children's reading performance in the class and tests.

Cognitive learning

Cognitive learning outcomes are measured by questions about atom, molecule and the modern periodic table. Children answered five multiple choice questions in the pre-test and got one point for each correct answer. They got zero points for each wrong answer. After interacting with the robot, children were asked to do an identical post-test. The cognitive learning score for each dependent variable was calculated by subtracting the post-test score with the pre-test score. Scoring for Cognitive Learning:

$$S(post) - S(pre) = CL$$

S: Score
CL: Cognitive Learning

Affective outcomes

Affective outcomes were measured by the Self-Assessment Manikin [34] (see Fig.7,8). SAM questionnaire has three dimensions, Pleasure, Arousal and Dominance. For pleasure, SAM ranges from a happy face to an unhappy face; for arousal, SAM ranges from an excited figure to a calm one; for dominance, SAM ranges from a very small figure representing a feeling of being submissive to a very large figure representing a powerful feeling [36]. During the experiment, children was a little bit confused by those words: pleasure, arousal and dominance, which was hard for them to understand. Therefore, the experimenter needed to explain how to choose the figure that represents their feelings and emotions for the participants. To be specific, the experimenter said, for the first row, the first one means that you are extremely happy now, and the last one means that you are extremely unhappy. For the second row, the first one means that you are extremely excited right now, and the last one means that you are extremely calm. For the third row, if you think you are controlled by the robot, you can choose the first one, and the last one means you think that are leading the robot to do this activity. It was easier for children to understand with the explanations and the vivid figures.

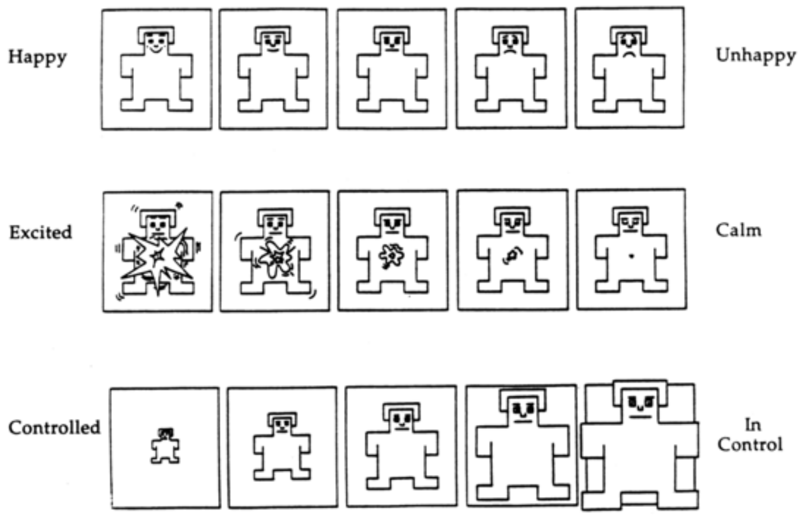

Fig. 7. The Self-Assessment Manikin.

Fig. 8. A participant who is doing the Self-Assessment Manikin survey.

Perceived competence
Perceived competence was measured by an adjective-based instrument developed by Peters et al [32]. It is assessed using 20 adjectives (translated from Dutch: *bossy, nagging, clumsy, friendly, popular, playful, follower, loner, angry, honest, fight, knowledgeable, boring, nice, listener, confident, educational, helpless, dumb*) on a three-point Likert scale. Children rated whether each word would describe the robot (yes, sometimes/maybe, no). This measure was used to investigate if children perceived the robot's role (tutor/learner) as high competent or low competent.

4.3 Adult experiment

An additional experiment was arranged because we found that the content was very difficult for the children to understand during the experiment. As a result, the pre and posttest may have limited information revealing their cognitive learning outcomes. Therefore, we recruited ten adults (we thought the content was easier accepted by adults) to do the same experiment to see their cognitive learning outcomes with the same robot setting.

Participants

Ten Chinese adults were invited to participant in this additional experiment, three males and seven females (aged 24 - 47 years old). Those participants all had no experience interacting with NAO robots. Seven of them reported that they would like to figure out things on their own, three of them would like to listen to the teacher first. Nine of them self-reported that they have good reading comprehension ability.

Procedures

The experiment was carried out in the same way as previous two. Further, participants were briefly interviewed after the experiment.

5. Results

5.1 Cognitive learning outcomes

To investigate the effect on cognitive learning outcome, we performed a Repeated Measures Analysis with independent variables: role (Learner and Tutor) and reading comprehension level (high and low), and cognitive learning outcomes (pre and post) as dependent variable. We found that children got higher scores after this activity (see Fig.9), and this effect was significant ($F(1,28) = 5.052$, $p = 0.033$). Firstly, roles have no significant effect on children's cognitive learning outcomes ($F(1, 28) = 0.101$, $p = 0.753$. Secondly, we noticed that children with higher reading compression ability got higher scores in this activity(M(pre)=2.2, M(post)=3.35), but children with low reading comprehension actually performed worse during this activity (M(pre)=2.7, M(post)=2.5) (see Fig.10), and this interaction effect of intervention and comprehension was significant ($F(1,28) = 4.850$, $p = 0.036$). Our results thus show that children learned from the activity in general, that this learning effect was moderated by reading comprehension, but that robot role did not have an influence on this.

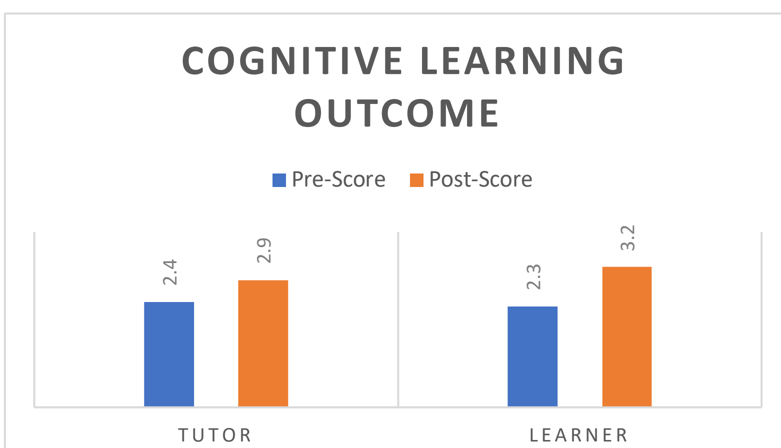

Fig. 9. Means of Tutor and Learner's pre and posttest scores ($F(1,28) = 5.052$, $p = 0.033$).

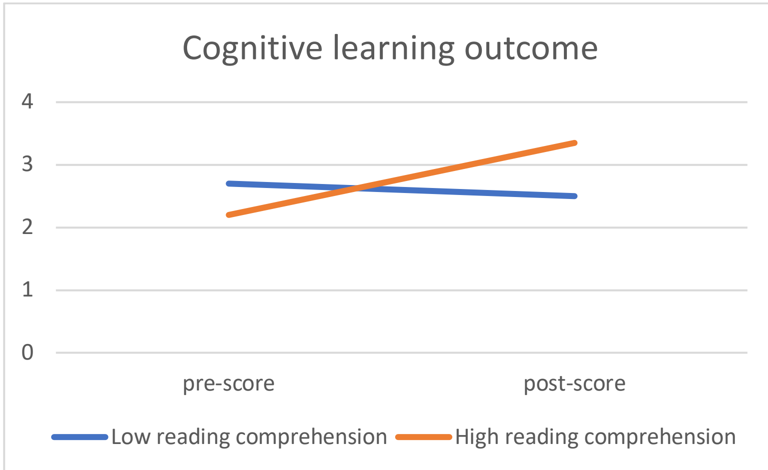

Fig. 10. Interaction effect between intervention and reading comprehension ability. High reading comprehension helps learning.

5.2 Affective learning outcomes

Then, we analyzed children's affective learning outcomes by measuring their Valence, Arousal and Dominance after they interacting with the robot (see Fig.11). We performed Multivariate Tests with independent variables role (Learner and Tutor), and affective learning outcomes as dependent variable. Firstly, we needed to recode Valence and Arousal because we wanted it to be in the order of negative to positive. Therefore, we used Compute Variable in SPSS, we used Numeric Expression: 9-Valeence and 9-Arousal. We found that roles have no significant influence on the perceived emotions (F (3, 26) =0.266, p=0.849) (Valence: F (1,28)=0.741, p=0.397, Dominance: F(1,28) = 0.325, p=0.573, Arousal: F(1,28)=0.072, p(Arousal) = 0.790). From the design of the questionnaire we know that "5" is a neutral score. And from the estimated marginal means we found that students are relatively neutral. But we do find that children felt a little bit dominant when the robot plays a learner role (MD (tutor) = 5.067, MD (learner) = 5.533), but they were not aroused after this activity (MA (tutor) = 4.600, MA (learner) = 4.800). However, they are higher in Valence when they interact with a tutor robot (MV (tutor) = 6.467, MV (learner) = 6.000). In all, role has no significant effect on the perceived emotion.

5.3 Perception

Last but not least, we performed a Multivariate Tests with independent variables role (Learner and Tutor), and the twenty adjectives as dependent variable to check whether these children perceived the robot roles differently. And we did Bonferroni correction for multiple comparisons. We found that children perceived *Clumsy* (F(1,28)=4.295, p=0.048), *Honest* (F(1,28)=16, p=0), *Confident* (F(1,28)=4.295, p=0.048), *Educational* (F(1,28)=6.236, p=0.019), *Helpful* (F(1,28)=4.846, p=0.36), *Dumb* (F(1,28)=5.814, p=0.023) significantly different. For other adjectives: *Bossy, Nagging, Friend, Popular, Playful, Follower, Loner, Angry, Fight, Knowledgeable, Boring, Nice, Listener, Helpless,* the result showed no significant difference.

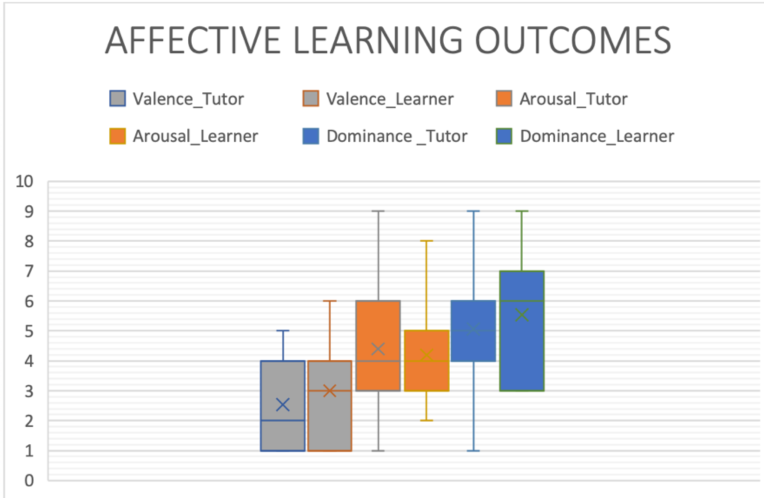

Fig. 11. Affective learning outcomes of Tutor and Learner.

5.4 Observations during the experiment

Children's behaviors when interacting with the robot reveals their emotions. We observed their learning behaviors as well as their confidence level, and took notes (see appendix 4) during this experiment and found that they showed some positive and negative emotions (See Fig.12).

Children's behaviors

We noticed that children stared at the robot when they entered the experiment room showing their curiosity. For instance, one child asked, "why this robot has only three fingers?". What's more, three of the participants even touched the robot's head and hands trying to make it do certain poses. Some extraverted children would ask a lot of questions about the robot with excitement and happiness.

Positive	**Negative**
Curious	Doubtful
Excited	Hesitation
Happy	Unconcentrated
Satisfied	Unhappy
Confident	Unsure
Decisive	Depressed
Entertained	Nervous
Motivated	Uncertain

Fig. 12. Positive and negative feelings of the children during the experiment.

Children's confidence level

Sometimes, children were confident and decisive to answer the robot's questions and were satisfied with the feedbacks. Some were even entertained by the robot's words and behaviors. When the robot encouraged them, some children were motivated and showed more confidence. However, some children showed that they were lack of confidence during the interaction. For example, they might be doubtful and uncertain if they didn't know the answer to the robot's question. They sometimes acted very slowly because of hesitation. Plus, some of the children turned to the experimenter for help if they were unsure about the answer. Plus, they were nervous to provide the robot with answers. Even some children felt unhappy and depressed when they were told to be wrong by the robot (One of the participants was even depressed and wanted to stop this learning activity).

Distraction

We also observed that some participants were unconcentrated because of two reasons. Firstly, they had slides, reading materials and the robot with information at the same time during the interaction. They sometimes were distracted by other materials other than the robot. As a result, they missed some behaviors and words expressing by the robot. Secondly, the external environment also affected their powers of concentration. Because we had this activity at a Chinese school, children had breaks between classes. They would run and shout during the break in the hallway. Some of our participants were distracted by those noise according to our observation.

5.5 Additional experiment

Ten adult participants all got higher scores after reading the material and interacting with the robot (see Fig.13). Seven of them got full post-scores after this activity. We performed a one-way repeated measure ANOVA with independent variable role, and cognitive learning outcomes as dependent variable. We found that adults significantly learned from this activity ($F (1,8) = 7.692^b$, $p=0.024$), but role had no effect on their learning outcomes ($F (1,8) = 1.231^b$, $p=0.299$), which is the same result as we got from the children in the previous study. Adults reported that they assumed that robots should be intelligent and know everything before this activity. However, they found that in this experiment, the robot could be lack of competence.

Fig. 13. Additional experiment with adults.

6. Discussion

In this experiment, the robot was designed to play two roles, tutor and learner, through verbal and non-verbal behaviors. For the tutor robot, it showed high competence by giving lectures and asking children questions. The robot checked whether the answer was right or wrong. If it was right, the robot praised, if not, the robot corrected the answer and gave encouragement. For the learner robot, it showed low competence by showing ignorance and seeking help from the children. The learner robot never judged children's answers, it just let the children explore with the help of reading materials. These two roles were designed based on human teachers' verbal behaviors. Non-verbal behaviors were based on previous child-robot interaction. According to Peters et, al., stable body posture and frequent hand gestures relate to high competence. Non-stable and less frequent hand gestures are perceived as low in competence [32]. In all, we designed the robot for this experiment with verbal and non-verbal behavior according to previous studies.

The results indicate that children did learn from this activity, and according to their different reading comprehension ability, their cognitive learning outcomes were different. Children were given the same reading materials, answer cards, and questionnaires. They were assigned to two different roles of the robot (tutor or learner), and the result shows that they did perceive robots differently based on their verbal and non-verbal behavior when it comes to *Clumsy*, *Honest*, *Confident*, *Educational*, *Helpful* and *Dumb*, which indicates that we designed the robot properly. However, to explain why they cannot perceive some important adjectives which describe competence such as *knowledgeable*, we have the following assumptions. Firstly, the sample is too small. Only 15 children for each group is far from enough to see whether it is significantly different. Secondly, according to our additional experiment with adults, people have stereotypes towards robots. They assume that robots should be smart and intelligent, and robots can be his/her teacher but not the other way around. Therefore, in future work, not only should we get more participants to take part in the experiment, but also, we should do more effort into making the role of learner even lower in competence. For example, for the robot design, we noticed that appearance would cause a first impression in people toward the robot. Perhaps a tutor robot should be taller and a learner robot should be smaller.

Children got different cognitive learning outcomes from this activity, which means eventually they learned (see Fig. 14). But the data indicates that they seem to learn from the reading materials. In our hypotheses, we want to see if a learner role results in higher learning outcomes compared to a tutor role if only for children with high reading comprehension ability and vice versa. It turns out that children with high reading comprehension ability did got better cognitive learning outcomes. However, this was not modified by the robot's role. A plausible explanation is therefore that they learned from reading the material, but not from the robot. It might because the content of this experiment was too difficult for these children to learn. So, they got nervous and frustrated learning things that they could not understand. Some children didn't pay much attention doing the pre and post-test because they didn't know how to answer those questions. Perhaps the fact that the robot did not correct the student in the learner role also hampered learning for the children with a low reading skill.

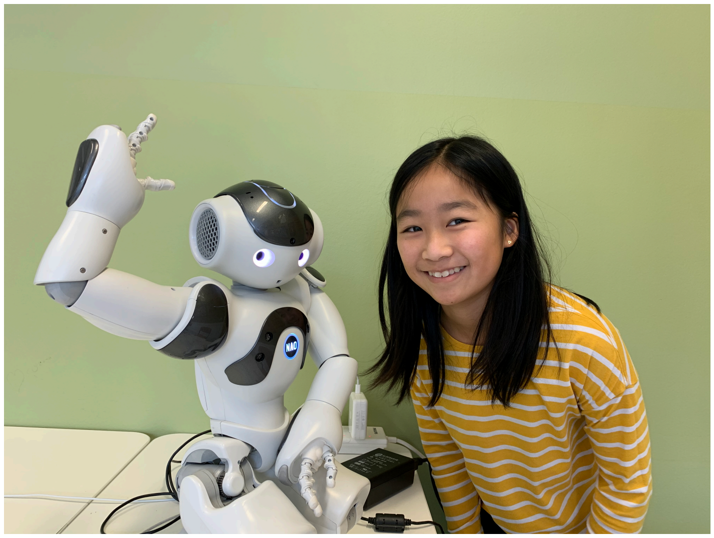

Fig. 14. A participant with the Nao robot.

Another possible reason might be the short-term character experiment. In this study, children only got about ten minutes to play with the robot, which made it hard for them to build a concrete relationship with the robot. Children could not perceive the robot's competence and they had limited time to learn from/ teach the robot. If we can design an experiment that lasting for two weeks, children would have more time to get familiar with the robot and interact with the it.

Concludingly, in future work, we need to get more participants, a lower the level of difficulty of the task and a design that better pronounces the difference between the two roles but still allows the children to learn the material, and a long-term experiment.

For the affective learning outcomes, we didn't see significant different between the tutor robot and the learner robot. Besides the reason that our sample is too small, we should pay attention to their culture background, they are Chinese Dutch. Most of them have at least one parent from China, but they grow up in the Netherlands, which means they are likely to live in a two-culture family. According to Hofstede's cultural dimensions theory [36] (see Fig.15), Chinese people are more restraint than people in the Netherlands, which means from the Chinese culture background point of view, children are taught to not showing their emotions too much. If you are extremely happy, just show a little. If you are very excited, don't say it out loud. Therefore, when we use the Self-Assessment Manikin Measurement, Chinese people would prefer to choose neutrally. It is possible that they hold their emotion, and choose some neutral options because of their Chinese culture background. In the future, we could improve this setting by encourage the children to show their true feelings. For example, we can leave them alone to do the survey without watching them to choose, and we can also let them know that being true to their feelings is the key to help the scientist to do the research. Last but not least, children may not know the exact meaning of Valence, Arousal and Dominant, which influences the accuracy of their choices.

Next time, researchers could give a short lecture in a classroom to all participants to help them understand the meaning of those words. Furthermore, research could use other method to measure participant's emotions. For example, they could use Geneva Emotion Wheel [40], which is a good method to help people describe their feelings.

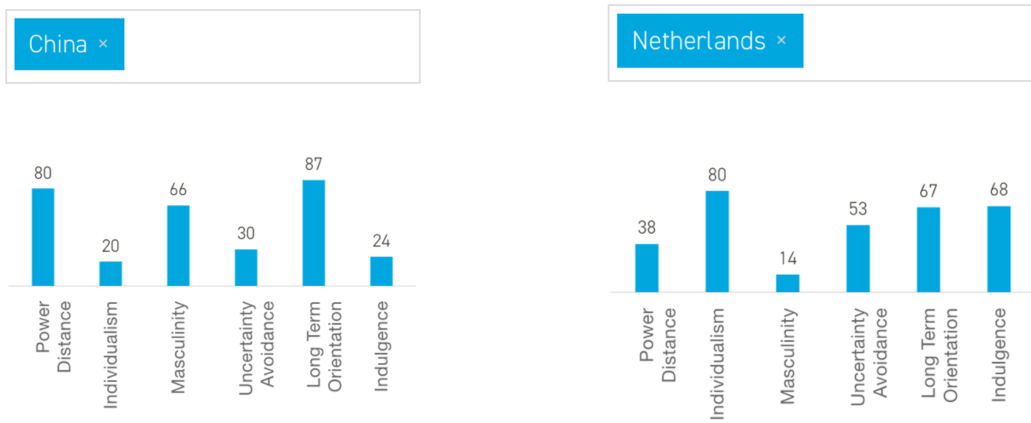

Fig. 15. China and Netherlands' Hofstede insights.

Fig. 16. Geneva Emotion Wheel (Scherer, 2005)

In this experiment, the different roles have no effect on learning outcome, and we did not find an interaction effect between role and reading comprehension. As such, we cannot conclude that there would be a benefit of adapting the role to the child based on reading comprehension.

7. Conclusion and further research

This work is an attempt to study the effect of robot competence expressed as a tutor versus a learner role on children's cognitive and affective outcomes and the additional effect of adaptive roles. We based the design of our roles on human being's verbal behaviors and non-verbal behaviors based on previous robot studies. In this experiment, the different roles have no effect on learning outcome, and we did not find an interaction effect between role and reading comprehension. As such, we cannot conclude that there would be a benefit of adapting the role to the child based on reading comprehension. It turns out that children did learn from this activity, but probably from the material and the extent to which the robot played a role is not certain. In this research, the bottom line is that we showed that children with different reading ability levels get different learning outcomes. Concludingly, in future work, we need to get more participants, a lower the level of difficulty of the task and a design that better pronounces the difference between the two roles but still allows the children to learn the material, and a long-term experiment. What's more, the non-verbal behavior design needs more research. For instance, how to express different roles, and how those roles are perceived needs to be researched and tested in the future. Last but not least, we need more participants. During this study, we were in a special period, the corona pandemic, which means getting participants to do research is extremely difficult. As a result, we only recruited thirty children and ten adults to do this experiment. We need more samples to get significant results in the future.

Reference

1. Mubin O, Stevens C J, Shahid S, et al. A review of the applicability of robots in education[J]. Journal of Technology in Education and Learning, 2013, 1(209-0015): 13.
2. Kennedy J, Baxter P, Belpaeme T. Comparing robot embodiments in a guided discovery learning interaction with children[J]. International Journal of Social Robotics, 2015, 7(2): 293-308.
3. Hood D, Lemaignan S, Dillenbourg P. When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting[C]//Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. 2015: 83-90.
4. Looije, Rosemarijn, Mark A. Neerincx, & Vincent de Lange. "Children's responses and opinion on three bots that motivate, educate and play." Journal of Physical Agents [Online], 2.2 (2008): 13-20. Web. 22 Feb. 2020.
5. Kanda T, Hirano T, Eaton D, et al. Interactive robots as social partners and peer tutors for children: A field trial[J]. Human–Computer Interaction, 2004, 19(1-2): 61-84.
6. Looije R, Neerincx M A, Lange V. Children's responses and opinion on three bots that motivate, educate and play[J]. 2008.
7. Tanaka F, Matsuzoe S. Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning[J]. Journal of Human-Robot Interaction, 2012, 1(1): 78-95.
8. Janssen, Joris B., et al. "Motivating children to learn arithmetic with an adaptive robot game." International Conference on Social Robotics. Springer, Berlin, Heidelberg, 2011.
9. Esubalew, T., et al. "A step towards developing adaptive robot-mediated intervention architecture (ARIA) for children with autism." IEEE Transactions on Neural Systems and Rehabilitation Engineering 21.2 (2012): 289-299.
10. Baxter, Paul, et al. "Robot education peers in a situated primary school study: Personalisation promotes child learning." PloS one 12.5 (2017).
11. Nourbakhsh I R, Bobenage J, Grange S, et al. An affective mobile robot educator with a full-time job[J]. Artificial intelligence, 1999, 114(1-2): 95-124.
12. Krathwohl D R, Anderson L W. A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives[M]. Longman, 2009.
13. Krathwohl D R, Bloom B S, Masia B B. Taxonomy of educational objectives: The classification of educational goals; Handbook II: Affective domain[M]. David McKay Company, Incorporated, 1956.
14. Krathwohl D R. A revision of Bloom's taxonomy: An overview[J]. Theory into practice, 2002, 41(4): 212-218.
15. Mubin O, Stevens C J, Shahid S, et al. A review of the applicability of robots in education[J]. Journal of Technology in Education and Learning, 2013, 1(209-0015): 13.
16. Belpaeme T, Kennedy J, Ramachandran A, et al. Social robots for education: A review[J]. Science robotics, 2018, 3(21): eaat5954.
17. Henkemans O A B, Bierman B P B, Janssen J, et al. Using a robot to personalise health education for children with diabetes type 1: A pilot study[J]. Patient education and counseling, 2013, 92(2): 174-181.
18. Hsieh E. "I am not a robot!" Interpreters' views of their roles in health care settings[J]. Qualitative Health Research, 2008, 18(10): 1367-1383.
19. Tanaka F, Matsuzoe S. Learning verbs by teaching a care-receiving robot by children: An experimental report[C]//2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 2012: 253-254
20. Okita S Y, Ng-Thow-Hing V, Sarvadevabhatla R. Learning together: ASIMO developing an interactive learning partnership with children[C]//RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication. IEEE, 2009: 1125-1130.
21. Hood D, Lemaignan S, Dillenbourg P. When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting[C]//Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. 2015: 83-90.
22. Kanfer R, Ackerman P L. Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition[J]. Journal of applied psychology, 1989, 74(4): 657.
23. Grasha A F. A matter of style: The teacher as expert, formal authority, personal model, facilitator, and delegator[J]. College teaching, 1994, 42(4): 142-149.
24. J Broekens, R Peters, F Kaptein, M Neerincx. (2019) DR 2.4: Integration and refinement.
25. Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. Science Robotics, 3(21).
26. den Brok P, Levy J, Brekelmans M, et al. The effect of teacher interpersonal behaviour on students' subject-specific motivation[J]. The Journal of Classroom Interaction, 2005: 20-33.
27. Parsons S A. Adaptive teaching in literacy instruction: Case studies of two teachers[J]. Journal of Literacy Research, 2012, 44(2): 149-170.
28. https://www.interactive-robotics.com/onderwijs/
29. Jamet F, Masson O, Jacquet B, et al. Learning by teaching with humanoid robot: a new powerful experimental tool to improve children's learning ability[J]. Journal of Robotics, 2018, 2018.
30. Burch G F, Heller N A, Burch J J, et al. Student engagement: Developing a conceptual framework and survey instrument[J]. Journal of Education for Business, 2015, 90(4): 224-229.

31. Cydney H Dupree and Susan Tufts Fiske. Universal dimensions of social signals: Warmth and competence. In Social Signal Processing, pages 23– 33. Cambridge University Press, 2017.
32. Peters R, Broekens J, Neerincx M A. Robots educate in style: The effect of context and non-verbal behaviour on children's perceptions of warmth and competence[C]//2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE, 2017: 449-455.
33. Sosniak L A. Bloom's taxonomy[M]. Chicago, IL: Univ. Chicago Press, 1994.
34. Bradley M M, Lang P J. Measuring emotion: the self-assessment manikin and the semantic differential[J]. Journal of behavior therapy and experimental psychiatry, 1994, 25(1): 49-59.
35. Morris J D. Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response[J]. Journal of advertising research, 1995, 35(6): 63-68.
36. Vitell, S.J., Nwachukwu, S.L. & Barnes, J.H. The effects of culture on ethical decision-making: An application of Hofstede's typology. J Bus Ethics 12, 753–760 (1993). https://doi.org/10.1007/BF00881307
37. Alemi, M., Meghdari, A., & Ghazisaedy, M. (2015). The impact of social robotics on L2 learners' anxiety and attitude in English vocabulary acquisition. International Journal of Social Robotics, 7(4), 523-535.
38. Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. Science robotics, 3(21).
39. Edwards, A., Edwards, C., Spence, P. R., Harris, C., & Gambino, A. (2016). Robots in the classroom: Differences in students' perceptions of credibility and learning between "teacher as robot" and "robot as teacher". Computers in Human Behavior, 65, 627-634.
40. Sacharin, V., Schlegel, K., & Scherer, K. R. (2012). Geneva emotion wheel rating study.