Learning English Words with the Aid of an Autonomous Care-Receiving Robot in a Children's Group Activity

Shizuko Matsuzoe[1], Hideaki Kuzuoka[2], and Fumihide Tanaka[3]

***Abstract*— We studied educational support delivered through a care-receiving robot (CRR) in a children's group activity intended to promote the learning of English words by teaching the robot. Our prior study investigated the feasibility of the CRR for providing educational support in a situation where a child played with the robot by him/herself. Our research uncovered several impactful effects of the CRR for enhancing childhood education. However, the results were not sufficient to confirm more practical contributions of the CRR toward learning. In this paper, we report on a field experiment we conducted with a group of children at a Japanese kindergarten aged 5–6 years. Our goal was to verify the feasibility of an autonomous CRR for facilitating the learning of English words in an educational setting that closely resembled reality. The experiment was conducted on a group of roughly seven children who participated in an animal gesture game with the robot for four days to learn six English words/names for animals. There were a total of 15 participants, and we held two experimental sessions. In order to compare the educational effects between learning with the aid of the CRR and with an expert robot, both robots were introduced concurrently into classrooms. The experimental results showed that the autonomous CRR was more effective in promoting English vocabulary acquisition among preschoolers compared to the expert robot.**

I. INTRODUCTION

In recent years, the use of personal robots in educational contexts has been attracting attention, with several studies having focused on their use as support tools for learning. Through our review of the current literature, we identified two types of educational support robots. The first assumes the role of the teacher or tutor [1], [2], [3], [4], [5]. Although these studies have demonstrated that the robots were educationally effective, there were still several concerns regarding its effectiveness with younger learners. For example, children would often become passive during lessons. Additionally, there was the issue of Robot Ethics [6]. Consequently, a different type of robot was proposed, one that could interact with children on their level, including being able to solve problems with them or play games. A care-receiving robot (CRR) [7], [8] is a typical example of this second type. It is designed to exhibit weak or incomplete behaviors in front of children. If the robot answers a question incorrectly, the children can teach it the correct response. Thus, the CRR promotes spontaneous learning by allowing children to teach it what they know. We focused on these advantages and conducted various field experiments at an English language school for Japanese children (3–6 years of age). Tanaka et al. demonstrated that the CRR can help children learn English verbs in an environment in which the children interact with the robot on a one-on-one basis [9]. Matsuzoe et al. investigated whether differences in how smartly robots behaved affected vocabulary acquisition in young English language learners; their findings indicated that an appropriate level of robot intelligence was instrumental in promoting vocabulary learning among children [10].

From these results, it gradually becomes evident that weakness or incompleteness demonstrated by a robot has the potential to enhance children's learning. Although we managed to verify the feasibility of the CRR for enhancing learning in settings where children played with the robot by themselves, we could not determine the practical utility of the CRR for children's educational support. Real educational situations are not limited to just one-to-one lessons. In the kindergarten and others, it is often the case that children attend various group activities. Therefore, we have expanded the previous experimental situation to include playing with the robot and multiple children simultaneously. We also need to investigate the effect of the continual use of the robot on children's learning. In addition, it would be important to verify if an autonomous robot can help with childhood education in a realistic context, such as in a group activity with multiple children.

To do so, we conducted a field experiment in classrooms at a kindergarten in Tsukuba. We introduced an animal gesture game using the robot, and had children attending the kindergarten participate in the experiment. We preprogramed an autonomous CRR to play the animal gesture game by allowing children to teach it the correct gestures. This weakness type robot was introduced concurrently with an expert type (which knew all the correct gestures from the start) to compare their educational impact.

The CRR is described in greater detail in Section II of this paper, while the field experiments we performed involving kindergarten children in a group language learning activity are explained in Section III. Section IV presents the results obtained with each robot, Section V discusses these results and possible limitations of the experiment, and Section VI concludes this paper.

[1]S. Matsuzoe is with the Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, 305-8573, Japan shi25tty at gmail.com

[2]H. Kuzuoka is with the Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, 305-8573, Japan kuzuoka at iit.tsukuba.ac.jp

[3]F. Tanaka is with the Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, 113-8656, Japan fumihide.tanaka at gmail.com

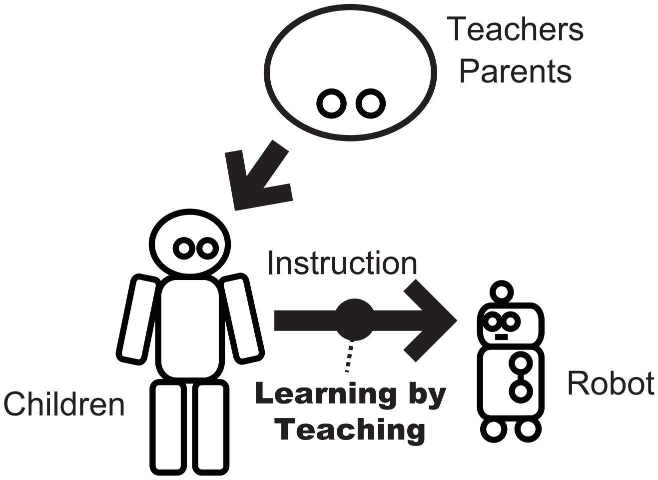

Fig. 1. Concept of care-receiving robot for supporting children's education [7], [8]

Fig. 2. Experimental field/setting

II. CARE-RECEIVING ROBOT (CRR)

A CRR is a robot that is designed to receive *care* from a human user. To achieve this, the robot shows weakness or incompleteness, such as falling down and answering questions incorrectly. Fig. 1 shows an example of how to use the CRR for children's educational support. First, the teachers/parents assign the children a learning topic, for example, pertaining to English, math, or manners). The CRR then makes a deliberate mistake, thereby prompting further instruction by the children. The children's own spontaneous teaching eventually enhances their learning.

In our prior study evaluating the feasibility of CRRs for providing educational support, we conducted several field experiments in which a child played with the robot face to face on a one-on-one basis. The results suggested that the deliberate mistakes made by the robot could promote the acquisition of new English words in children. However, formal learning typically includes multiple children, such as in a classroom or group activity. Therefore, our results were insufficient to determine the full extent of its possibilities for facilitating children's learning. The next phase of our investigation will need to address the effectiveness of the CRR as an educational support tool within a realistic classroom setting in which there are multiple children.

III. A FIELD EXPERIMENT INVOLVING A KINDERGARTEN GROUP ACTIVITY

A. Goals of the experiment

Our goal was to verify the feasibility of the CRR as an educational support tool within a learning environment made to resemble reality as much as possible. We aimed to answer the following research questions:

(1) Can the CRR promote children's acquisition of new English words?
(2) Can the CRR engage children's interest in the animal gesture game?

Below is an outline detailing the environments/settings for the experiment:

- The robot is introduced into a group activity at the kindergarten.
- Use of the autonomous CRR.
- Children participate in the experiment with the robot for four days.

We had children participate in an animal gesture game with the robots to learn English words/names for animals. Assuming that the CRR had similar effects on learning as those we had observed in our previous study, the CRR could be expected to facilitate the acquisition of new English words in the whole group. Moreover, there was the probability that the robot could engage children's interest in the animal gesture game. In order to compare the educational effects between learning with the aid of the CRR and with an expert robot, both robots were introduced concurrently into classrooms.

B. Experimental field

The experiment was conducted in actual classrooms with the cooperation of teachers at a kindergarten in Tsukuba. Fig. 2 shows an example of the classrooms in which the field experiments were conducted. We left toys in each classroom during the entire experiment and allowed the children to play with them alongside the robots to maintain an environment that was typical and familiar to them. The robots used in this experiment were sitting in the center of the room, and an appropriate distance was maintained between them. Two camcorders were set up in the corner of the room to record the behaviors of the children and the robots.

There were a total of 15 participants (children 5–6 years of age; ten girls and five boys) in the experiments, who regularly spoke Japanese. They were divided into two groups, with which we held two experimental sessions. This experiment was approved by the Ethical Committee of the University of Tsukuba. We explained our research goals and experiment to the parents of the children. We began the experiment after receiving written consent from each parent.

Fig. 3. Direct-teaching: children help NAO step-by-step by guiding its hand, and teach it how to gesture an elephant.

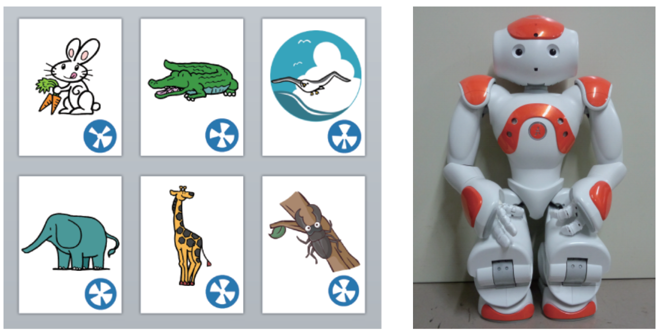

Fig. 4. Equipment used: (left) animal cards with specific landmarks (Naomarks); (right) Aldebaran Robotics' NAO

C. Experimental task and equipment

In this experiment, we asked the participants to play an animal gesture game with the robots. When the children showed an animal card, the robot recognized the animal illustrated on the card. Then, the robot attempted to make an animal gesture corresponding to the card shown. When the robot demonstrated an incorrect gesture, the children were able to take the robot by the hand and teach it a gesture in a stepwise manner (direct-teaching: Fig. 3). If the children succeeded in teaching the robot the correct animal gesture, it would demonstrate the corresponding animal gesture. Six animal gestures were targeted in this game: those of a rabbit, elephant, alligator, giraffe, sea gull, and stag beetle. The animal cards used for the game are shown in Fig. 4 (far-left panel). These animals were randomly divided into two groups, with consideration for counterbalancing the difficulty level of the English words, which was determined according to their frequency of use by native speakers. Each three-animal group of words and gestures was assigned to one robot, so that it could play with the children.

Aldebaran Robotics' humanoid robot, NAO, was used for this experiment (Fig. 4, far-right panel). We implemented an autonomous system within its behavioral controls that would enable the NAO to play the animal gesture game. The basic specifications of the autonomous robot we used for the animal gesture game are as follows:

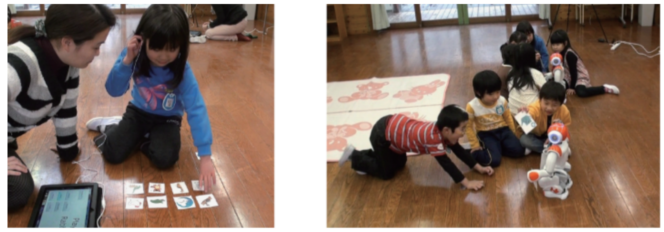

Fig. 5. Scenes from the experiments: (left) pre-test; (right) free-play time

- Robust recognition of the animal cards presented by participants using Naomarks, which are special landmarks with specific patterns designed for NAO (provided by Aldebaran Robotics) that are rotation and scale invariant.
- Real-time determination of the animal mimicked using the robot's motion via children's direct-teaching; the autonomous system distinguishes an animal gesture based on the variation of the joint angles, which move characteristically with respect to each animal.

We used small animal cards, a tablet PC, and earphones to conduct the pre/post-tests. Fig. 5 (far-left panel) shows a photograph of a participant in the pre-test using these tools. Each participant was able to hear the sound of a native speaker pronouncing the specific English word/name of the animal from the tablet PC.

D. Experimental design

We used the following two types of robots in the experiment:

- **Expert robot**: the robot got all of the animal gestures correct from the beginning.
- **CRR**: at the beginning, the robot could not demonstrate any animal gesture. When the children had taught it how to gesture for a certain animal, the robot was then able to demonstrate the animal gesture correctly.

In this experiment, we applied parallel comparison of expert robot and CRR. Our trial experiences had indicated that it would be difficult to differentiate whether children were engaged by the robot's novelty or failed attempt at gesturing if we used the CRR alone. Thus, we introduced these robots concurrently into the classrooms to verify the educational impact of the CRR and the expert type in an actual learning environment. We anticipated that children would choose the robot that they wanted to play with based on a comparison between the robots, displaying simple and natural choice behaviors.

E. Procedure

Fig. 6 depicts the experimental procedure. Children participated in the animal gesture game for four days. The following sections detail each stage of the entire procedure.

Pre-/Post-test: We conducted pre- and post-tests before and after playing the gesture game to investigate the degree of participants' acquisition of new English animal words/names. During this stage, each participant played a card game in which they heard native English pronunciations of a specific animal word/name and chose a corresponding

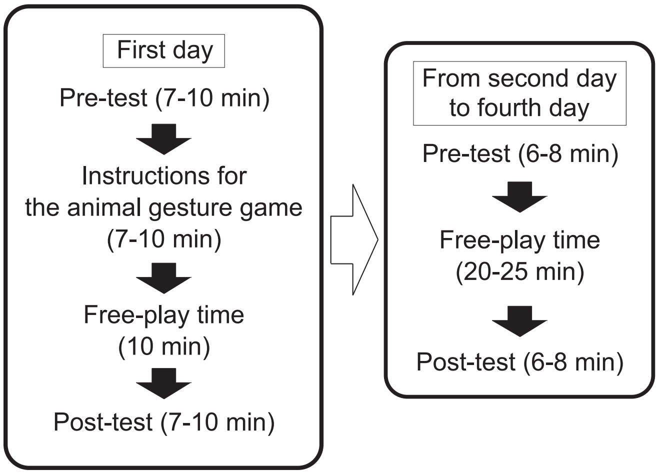

Fig. 6. Procedural flow of the experiment

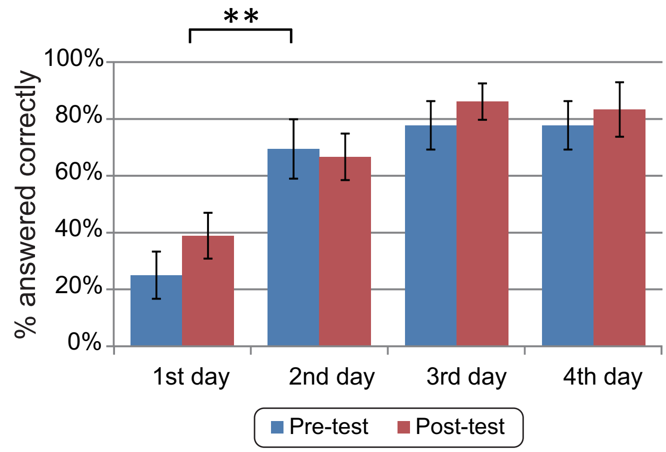

Fig. 7. Pre- and post-test results (with expert robot): each bar represents the average percentage of test questions answered correctly. we conducted the Wilcoxon signed-rank test, no significant differences between pre- and post-test scores were found on any day. However, the results revealed a significant difference between the post-test scores of the first day and pre-test scores of the second day [$Z(12) = -2.754,\ p < 0.01$].

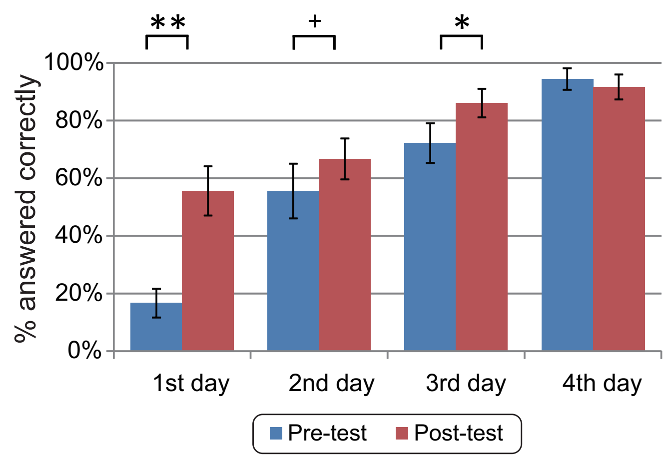

Fig. 8. Pre- and post-test results (with CRR): each bar represents the average percentage of test questions answered correctly. The Wilcoxon signed-rank test was conducted to ascertain if there were significant differences between pre- and post-test conditions. The results indicated significant differences between pre- and post-tests on the first day [$Z(12) = -2.625,\ p < 0.01$] and third day [$Z(12) = -1.933,\ p < 0.05$,]. On the second day, the difference between pre- and post-tests was marginally significant [$Z(12) = -1.823,\ p = 0.063$].

animal card. The experimenter presented eight cards in succession to each participant, and gave a response that included the Japanese name of the corresponding animal, such as "A rabbit is *usagi*," regardless of whether the participant's initial answer had been correct. There were six questions, one corresponding to each of the target animals for this game.

Instructions for the animal gesture game: This process was implemented on the first day only. The experimenter gave the participants instructions regarding the six animal gestures used for this game, demonstrating how to show a card to the robot and teach the robot an animal gesture. Then, the experimenter asked the children to show the card and teach the motion to the robot.

Free-play time: Fig. 5 (far-right panel) shows a photograph of participants during free-play. The experimenter asked the participants to freely play the animal gesture game among themselves. On the first day only, we switched the robots among the participants in the middle of their free-play so that they could experience playing with both robots. We permitted experimenters to provide minimal assistance to the participants when they encountered difficulties, such as an inability to show a card well or teach a desired gesture to the robot. If none of the above occurred, the experimenters were devoted to watching out for the safety of the participants during their play.

IV. RESULTS

A. Can the CRR promote children's acquisition of new English words?

Fig. 7 shows the average percentages of pre- and post-test questions answered correctly for three animal words/names allocated to the expert robot, while Fig. 8 shows these percentages for the three animal words/names allocated to the CRR. Both scores showed a gradual increase from one day to the next. However, we observed a marked increase in scores between pre- and post-tests on the first day for the CRR. The saturation score of the CRR learning situation was also slightly higher than that of the expert robot educational experience.

Additionally, we conducted a Wilcoxon signed-rank test, which revealed significant differences between pre- and post-test scores for the CRR on the first day [$Z(12) = -2.625,\ p < 0.01$] and third day [$Z(12) = -1.933,\ p < 0.05$]. On the second day, the difference between pre- and post-tests was marginally significant ($Z(12) = -1.823$, p=0.063). On the other hand, for the expert robot, no significant differences between pre- and post-test scores were found on any day. However, we had not anticipated that there would be a significant difference between post-test scores on only the first day and pre-test scores on the second ($Z(12) = -2.754,\ p < 0.01$).

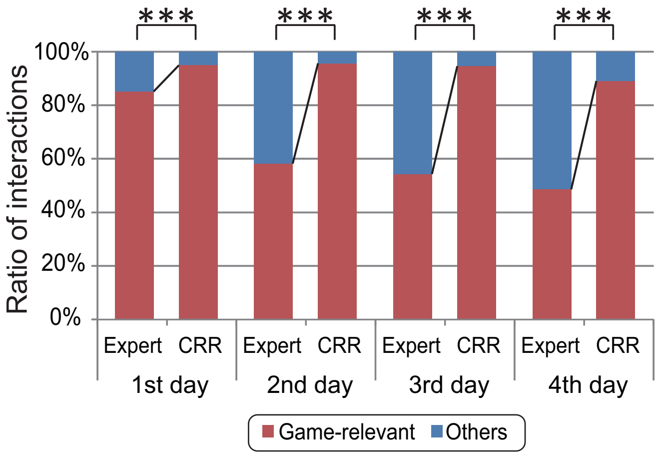

Fig. 9. Ratio of game-relevant participant interactions and other interactions: each color on a bar represents the percentage of game-relevant participant interactions and other interactions. The Chi-square test conducted showed that there were significant differences between interactions with the expert robot and CRR [first day: $\chi^2(1) = 11.767, p < 0.001$; second day: $\chi^2(1) = 59.898, p < 0.001$; third day: $\chi^2(1) = 76.735, p < 0.001$; fourth day: $\chi^2(1) = 63.865, p < 0.001$].

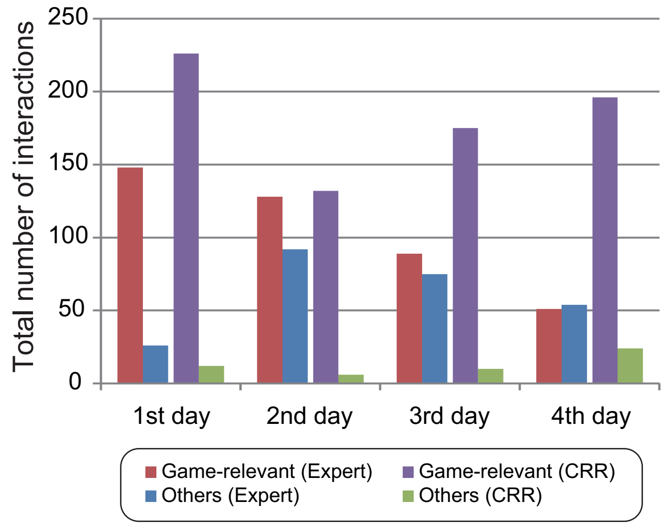

Fig. 10. Total number of game-relevant participant interactions and others: each bar represents the number of participant interactions with each robot by the day. The expert robot gradually decreases game-relevant participant interactions from the first to fourth days (red bars). Although game-relevant interactions decreased once on the second day with the autonomous CRR, it tended to maintain game-relevant interactions over the four-day period of our experiment (purple bars).

B. Can the CRR engage children's interest in the animal gesture game?

We conducted a behavioral analysis of the participants' interactions with each robot during the free-play time. The participants' interactions were divided into two categories: the first category included *game-relevant* interactions, such as showing an animal card to the robot, teaching the robot how to gesture for an animal, and praising the robot. We categorized other interactions (unrelated to the animal gesture game) as *others*, such as dressing the robot, hitting the robot needlessly, and giving the robot irrelevant objects. Behavioral video coding was also conducted by a third party according to these categories of behavior. The inter-observer reliability calculated from among three independent external coders was 0.91.

Fig. 9 shows the ratio of game-relevant participant interactions with the robots and other interactions. With the CRR, game-relevant interactions were prompted more frequently than were other interactions from the first day to the fourth. We conducted Chi-square tests, and the results showed that there were significant differences between interactions with the expert robot and CRR [first day: $\chi^2(1) = 11.767, p < 0.001$; second day: $\chi^2(1) = 59.898, p < 0.001$; third day: $\chi^2(1) = 76.735, p < 0.001$; firth day: $\chi^2(1) = 63.865, p < 0.001$].

Fig. 10 focuses on the total number of game-relevant participant interactions. We observed that the expert robot gradually decreased game-relevant interactions from the first day to the fourth. In contrast, although such interactions decreased once on the second day, the autonomous CRR tended to maintain these (game-relevant interactions) across all four days. Other kinds of interactions occurred as frequently as or less than 25 times throughout the entire experiment.

V. DISCUSSION AND LIMITATIONS

With regard to the acquisition of new English words, the results presented in Section IV-A indicate that the autonomous CRR was able to increase children's test scores gradually throughout the period of four days. We assume that the autonomous CRR in the group activity contributed toward reinforcing children's learning of English animal words/names. Regarding engagement of children's interest in the animal gesture game, we observed that more children played the animal gesture game with the autonomous CRR than with the expert robot throughout the duration of our experiment. Therefore, we suggest that the autonomous CRR could also engage participants' interest more than might the expert robot.

However, we did not expect the marked gap we observed between the post-test scores of the first day and pre-test scores of the second day. Upon closer inspection, we noticed that English animal words/names with low difficulty levels very often generated correct answers among our study participants. As we explained earlier, we determined the level of difficulty of each word based on its frequency of use among native speakers. For that reason, we suppose that Japanese children also get many opportunities to hear these low difficulty words in daily life as compared to words of a higher level.

This field experiment had several limitations. It should be noted that a rigorous control experiment was not conducted intentionally. The actual experimental field, the kindergarten, contained many factors/variables that affected children-robot interactions, such as friendships between the children, individual differences, and their physical or mental states at the time. We were unable to perform a rigorous control experiment in relation to these factors because of the time constraints for the field experiment. Therefore, the interdependence among these factors and effects of the order/sequence in which the children participated in the experiment could not be ascertained, among others.

VI. CONCLUSIONS

In this paper, we described a field experiment we conducted involving a parallel comparison of two autonomous robots. To verify the feasibility of autonomous CRRs for facilitating learning in group activities, the experiment was set in classrooms in an actual kindergarten setting, where children participated in an animal gesture game over four days with two robots and multiple classmates. We found from the results of our investigation, that the autonomous CRR successfully engaged the children and promoted the learning of English animal words/names. Our research on the effectiveness of CRRs in enhancing childhood education could expand to include other practical applications and contexts, such as in activities with smaller or larger groups and other educational support functions. In this experiment, we focused on only the wrong/incorrect behaviors of the CRR as weaknesses. There is the need to explore other effective weak behaviors of the CRR for facilitating children's learning in the future.

ACKNOWLEDGMENT

We would like to acknowledge the support we received from Higashi Kindergarten in Tsukuba, Japan, and from the KAKENHI (23680020) and JST PRESTO programs. We also thank the children and their parents, as well as the students of University of Tsukuba for their assistance with the project.

REFERENCES

[1] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "Interactive robots as social partners and peer tutors for children: a field trial," *Human-Computer Interaction*, vol. 19, no. 1, pp. 61–84, 2004.

[2] Z.-J. You, C.-Y. Shen, C.-W. Chang, B.-J. Liu, and G.-D. Chen, "A robot as a teaching assistant in an english class," in *Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies (ICALT 2006)*, 2006, pp. 87–91.

[3] T. Kanda, R. Sato, N. Saiwaki, and H. Ishiguro, "A two-month field trial in an elementary school for long-term human-robot interaction," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 962–971, 2007.

[4] J. Han, M. Jo, V. Jones, and J. H. Jo, "Comparative study on the educational use of home robots for children," *Journal of Information Processing Systems*, vol. 4, no. 4, pp. 159–168, 2008.

[5] M. Saerbeck, T. Schut, C. Bartneck, and M. D. Janse, "Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor," in *Proceedings of the 28th international conference on Human factors in computing systems(CHI 2010)*, 2010, pp. 1613–1622.

[6] N. E. Sharkey, "The ethical frontiers of robotics," *Science*, vol. 322, no. 5909, pp. 1800–1801, 2008.

[7] F. Tanaka and T. Kimura, "The use of robots in early education: a scenario based on ethical consideration," in *Proceedings of the 18th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2009)*, 2009, pp. 558–560.

[8] F. Tanaka and T. Kimura, "Care-receiving robot as a tool of teachers in child education," *Interaction Studies*, vol. 11, no. 2, pp. 263–268, 2010.

[9] F. Tanaka and S. Matsuzoe, "Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning," *Journal of Human-Robot Interaction*, vol. 1, no. 1, pp. 78–95, 2012.

[10] S. Matsuzoe and F. Tanaka, "How smartly should robots behave? : Comparative investigation on the learning ability of a care-receiving robot," in *Proceedings of the 21th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2012)*, 2012, pp. 339–344.