Children's anthropomorphism of second language robot tutors

Mirjam de Haas[1*], Rianne van den Berghe[2], Ora Oudgenoeg-Paz[2], Paul Leseman[2], Emiel Krahmer[1] Josje Verhagen[2], Paul Vogt[1], Bram Willemsen[1], Jan de Wit[1]

1 Tilburg center for Cognition and Communication (TiCC), Tilburg University, Tilburg, The Netherlands
2 Department of Special Education, Utrecht University, Utrecht, the Netherlands

*** mirjam.dehaas@tilburguniversity.edu** [1]

Abstract

This paper describes a study in which we investigated (1) whether the degree to which children anthropomorphize a robot changes after having interacted with it, and (2) whether there is a relation between children's anthropomorphism and their second language learning. We found that changes in children's perception of the robot was correlated with a higher learning gain. Children who attributed more human-like qualities after the tutoring sessions than before learned more than children who did not change their perceptions, or who decreased in the degree to which they anthropomorphized the robot. Our results highlight the need to consider anthropomorphism in designing robot-assisted education and adaptive tutoring systems.

Introduction

As robots are increasingly used for educational purposes, there is a growing need to understand the mechanisms of child-robot interactions. Robots have the potential to teach children in personalized interactions. Personalizing tutoring interactions can lead to higher learning gains [7]. Such interactions are often personalized regarding children's speed of learning, but their perception of and personal bond with the robot may also play a role. People have expectations of the robot's behavior and abilities, even before they interact with the robot [5]. When interacting with a social robot, people have a tendency to attribute human form, characteristics, and/or behaviors to the robot. This phenomenon is called anthropomorphism [1]. Children are particularly likely to assign cognitive and affective beliefs to robots, such as the ability to remember people and understand people's feelings [2].

The degree to which learners anthropomorphize robots may play an important role, as learning is first and foremost a social process [11]. Currently, it is not yet clear whether the degree to which learners anthropomorphize a robot affects how much they learn from it. Children who anthropomorphize the robot more, might interact with the robot similarly as they would interact with peers. This could be a benefit for learning, as literature on peer learning shows possible benefits of peers on learning [6, 8, 9, 12]. Robots may have similar benefits for learning when being treated as a peer. However, a

[1]The first two authors had an equal contribution to this paper.

certain level of anthropomorphism may be required to induce peer benefits. It is possible that a robot's benefit, and thus the child's learning in robot-assisted education, depends on the degree to which the learner anthropomorphizes it. This begs the question whether anthropomorphism and learning are related to each other in robot-assisted education. This has not been investigated yet to the best of our knowledge and is the central research question of our paper.

Research that comes closest is that of Chandra et al. [3], who investigated whether children's perception of a robot, in terms of intelligence, likability, and friendliness, affects children's learning in a learning-by-teaching paradigm. Twenty-five seven to nine-year-old children taught a NAO robot to write over the course of four sessions. There were two conditions: (1) the robot improved its handwriting, or (2) the robot did not improve its writing. Children in the first condition were able to perceive the robot's improvement by the last session, but this did not affect the robot's perceived intelligence, likability, and friendliness. However, their learning was correlated with the likability of the robot. In the second condition, children's learning was correlated with the friendliness of the robot. These findings need to be interpreted with caution due to the small sample size, but suggest that children's perception of the robot may be related to their learning. This paper also demonstrates the importance of multi-session studies as most results were found in the final interactions with the robot.

Our study is aimed at expanding previous work in two ways. First, it includes a larger sample. Second, it measures the degree to which children anthropomorphize the robot both before and after having interacted intensively with it, by assessing whether they perceive the robot more as a machine or more as a human. Our goal is to investigate (1) whether children's perception of the robot changes over time, and (2) whether there is a relation between children's anthropomorphism and their second language (L2) learning.

Materials and Methods

In this study, 119 Dutch children (58 boys) with an average age of 5 years and 8 months ($SD = 5$ months) participated in seven L2 (English) tutoring sessions with a SoftBank NAO robot and a tablet. In the iconic gesture condition (62 children), the robot acted out the word whenever the robot said one of the target words and deictic (pointing) gestures. For example, each time the robot said *big* it also moved the hands far apart from each other, during the rest of the interaction the robot sometimes pointed at the tablet to direct the child's attention to the tablet. In the no iconic gesture condition (57 children) the robot used only deictic gestures. Thirteen children dropped out and weren't included in the analyses. The original experiment included two more conditions, however for this paper we only focus on the two robot tutoring conditions. In seven robot tutoring sessions children were taught 34 different L2 words.

In this paper we focus on two different measures, each administered twice: 1) A perception questionnaire measuring the degree to which children anthropomorphized the robot, administered once prior to the very first tutoring session and once after the seventh and last session. 2) A comprehension test measuring children's L2 vocabulary learning gain, administered during an immediate post-test (one or two days after the last session) and during a delayed post-test (two to four weeks after the immediate post-test). Other measurements and a more detailed description of the tutoring sessions and experimental conditions can be found in Vogt et al. [10].

Perception questionnaire This questionnaire measured to what extent the children perceived the robot as a human or as a machine. It consisted of 12 questions. Each question could be answered with 'yes'/'no'/'I don't know' and had an open-ended query

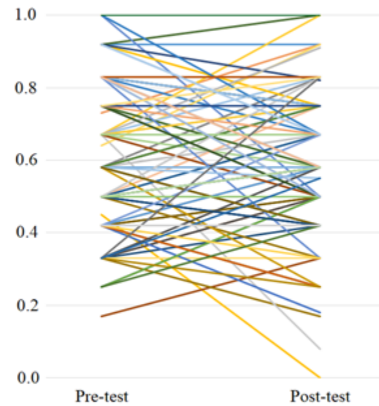

Figure 1. Individual trajectories from pre- to post-test anthropomorphism (proportion). The dashed line represents the group mean.

why children answered with this response. The items were based on Jipson and Gelman [4] in which was investigated to what degree children make a distinction between living and non-living items. The children were allotted one point for each 'yes' answer and their anthropomorphism score was the proportion of 'yes' answers. Thus, when a child answered all questions with a 'yes', this child would get a score of one and would consider the robot as a human rather than a machine.

Comprehension task This task measured to what extent children had learned the target words. The task showed three pictures (or videos for verbs) on the screen from which children had to choose the picture or video corresponding to the target word. Each target word was tested three times, to overcome chance level performance due to guessing. We did not test all 48 target words three times, as this would result in a test that is too long for such young children. We made a pseudo-random selection of 18 target words (see for a more detailed explanation [10]). Children were allotted one point when they chose the correct word at least two times (out of three).

Results

First, we looked at the perception scores of the children's pre- and post-test. A paired samples t-test did not show significant differences between children's pre- and post-test scores, $t(103) = 1.53, p = .130, d = .10$. However, as Figure 1 shows, there was large variability among the children in whether and how their perceptions changed between the pre- and post-test: from each level of pre-test anthropomorphism, the children's anthropomorphism either increased, decreased, or stayed consistent. We also explored whether children perceived the robot differently in the iconic-gesture condition compared to the condition with deictic gestures, using an ANOVA. We did not find any differences ($F(1, 102) = .00, p = .957, \eta p^2 = .00$).

Table 1 displays the correlation matrix between the children's comprehension and anthropomorphism. We found two significant correlations between children's anthropomorphism and learning, both with a small effect. Pre-test anthropomorphism was related to comprehension scores on the immediate post-test, $r(104) = -.208, p = .034$. The relation was negative, suggesting that children who anthropomorphized the robot more prior to starting the lesson series learned less than children who anthropomorphized the robot less. Children's change in

Table 1. Correlation Matrix Perception and Learning

	Comprehension	
Anthropomorphism	Immediate post-test	Delayed post-test
Pre-test	-.208*	-.137
Post-test	-.152	.094
Change	.036	.212*

* Significant at the .05 level.

anthropomorphism was related to comprehension scores on the delayed post-test, $r(106) = .212, p = .031$. This indicated that children who increased in the degree to which they anthropomorphized the robot learned more than children who did not increase in anthropomorphism.

Discussion

The goal of the current study was to investigate (1) whether children's perception of the robot changes over time, and (2) whether there is a relation between children's anthropomorphism and their L2 learning. The data showed that there are large individual differences in the degree to which children perceive the robot, in the way their perception changes over time, and that the change of their perception correlates with their L2 comprehension. Why this is the case is still unclear. As a next step the open-ended answers might give some insight on how the children's perception changed.

A possible reason for the relations between anthropomorphism and learning is that children who tended to anthropomorphize the robot prior to the lessons may have had high expectations of the robot's abilities, were disappointed with its performance during the lessons, and therefore learned less. Alternatively, these children may have struggled with learning and were disappointed with the help of the robot. Children who increased in anthropomorphism may have developed a bond with the robot that benefited their learning or the other way around, the children learned more and therefore developed a bond with the robot.

Future work can further investigate the role of children's expectations in robot-assisted education and explore adaptation to individual children's expectations and preferences. For example, for children who have high expectations the robot can adapt itself to act more human-like to maintain the children's expectations. It is still an open question whether managing children's expectations prevents them to decrease in anthropomorphism and whether this will benefit their learning.

Acknowledgment

This work has been supported by the EU H2020 L2TOR project (grant 688014). We would like to thank the children, their parents, and the schools for their participation. Furthermore, we would like to thank Laurette Gerts, Annabella Hermans, Esmee Kramer, Madee Kruijt, Marije Merckens, David Mogendorff, Sam Muntjewerf, Reinjet Oostdijk, Laura Pijpers, Chani Savelberg, Robin Sonders, Sirkka van Straalen, Sabine Verdult, Esmee Verheem, Pieter Wolfert, Hugo Zijlstra and Michelle Zomers for their help in collecting data.

References

1. C. Bartneck, D. Kulic, E. Croft, and S. Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1:71–81, 2009.

2. T. N. Beran, A. Ramirez-Serrano, R. Kuzyk, M. Fior, and S. Nugent. Understanding how children understand robots: Perceived animism in child–robot interaction. *International Journal of Human-Computer Studies*, 69(7-8):539–550, 2011.

3. S. Chandra, R. Paradeda, H. Yin, P. Dillenbourg, R. Prada, and A. Paiva. Do children perceive whether a robotic peer is learning or not? In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 41–49. ACM, 2018.

4. J. L. Jipson and S. A. Gelman. Robots and rodents: Children's inferences about living and nonliving kinds. *Child development*, 78(6):1675–1688, 2007.

5. S. Kiesler and J. Goetz. Mental models of robotic assistants. pages 576–577, 01 2002.

6. A. King. Transactive peer tutoring: Distributing cognition and metacognition. *Educational Psychology Review*, 10:57–74, 1998.

7. D. Leyzberg, S. Spaulding, and B. Scassellati. Personalizing robot tutors to individuals' learning differences. *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 423–430, 2014.

8. A. M. O'Donnell and A. King. *Cognitive perspectives on peer learning.* Routledge: New York and London, 1999.

9. K. Topping, S. Hill, A. McKaig, C. Rogers, N. Rushi, and D. Young. Paired reciprocal peer tutoring in undergraduate economics. *Innovations in Education and Training International*, 34:96–113, 1997.

10. P. Vogt, R. van den Berghe, M. de Haas, L. Hoffmann, J. Kanero, E. Mamus, J.-M. Montanier, C. Oranç, O. Oudgenoeg-Paz, D. H. García, F. Papadopoulos, T. Schodde, J. Verhagen, C. D. Wallbridge, B. Willemsen, J. de Wit, T. Belpaeme, T. Göksun, S. Kopp, E. Krahmer, A. C. Küntay, P. Leseman, and A. K. Pandey. Second language tutoring using social robots: A large-scale study. To appear in *the 2019 ACM/IEEE International Conference on Human-Robot Interaction*, 2019.

11. L. Vygotsky. Interaction between learning and development. *Readings on the development of children*, 23(3):34–41, 1978.

12. F. Yarrow and K. J. Topping. Collaborative writing: The effects of metacognitive prompting and structured peer. *British Journal of Educational Psychology*, 71:261–282, 2001.