Exploring the Role of Perspective Taking in Educational Child-Robot Interaction

Elmira Yadollahi[1,2]([✉]), Marta Couto[1], Wafa Johal[3], Pierre Dillenbourg[2],
and Ana Paiva[1]

[1] Group on AI for People and the Society (GAIPS),
INESC-ID & Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
`marta.couto@gaips.inesc-id.pt, ana.paiva@inesc-id.pt`
[2] Computer-Human Interaction in Learning
and Instruction Laboratory (CHILI), EPFL, Lausanne, Switzerland
`{elmira.yadollahi,pierre.dillenbourg}@epfl.ch`
[3] University of New South Wales, Sydney, NSW, Australia
`wafa.johal@unsw.edu.au`

Abstract. Perspective taking is an important skill to have and learn, which can be applied in many different domains and disciplines. While the ability to recognize other's perspective develops in humans from childhood and solidifies during school years, it needs to be developed in robotic and artificial agents' cognitive framework. In our quest to develop a cognitive model of perspective taking for agents and robots in educational contexts, we designed a task that requires the players (e.g., child and robot) to take the perspective of another, in order to complete and win the task successfully. In a preliminary study to test the system, we were able to evaluate children's performance over four different age groups by focusing on their performance during the interaction with the robot. By analyzing children's performance, we were able to make some assumptions about children's understanding of the game and select the appropriate age group to participate in the main study.

Keywords: Child-robot interaction · Spatial perspective taking · Children · Education · Gamification

1 Introduction and Background

The introduction of robots into education and interaction with children can revolutionize education as we know it. To have robots with capabilities to carry out educational roles, play games, be peers in the activities of a classroom, and at the same time, support learning in different forms is a challenging task. To achieve that we need to equip our robots with cognitive abilities that help them to become true learning companions. To endow the robots with cognitive abilities, we can either focus on the cognitive development, or the interaction capabilities of the robot, or develop both aspects simultaneously. One of the crucial aspects of educational scenarios is maintaining mutual understanding between the child

and the robot. To maintain such an understanding, it is inevitable for the child and the robot to develop a model of each other's mind and perspective.

Developmental psychology defines *Perceptual Perspective Taking* as understanding what other people see and their spatial or visual relationship with the objects in the environment. Taking others visual and spatial perspectives, consists of two levels that correspond to different developmental ages, the extent of perception, and their underlying mechanisms [10,20,21]. "Level 1" develops at around 24 months and corresponds to the ability to judge if an object is visible to another person (visual) or if it is positioned in their front or back (spatial) [14,18,22]. "Level 2" develops from 3–5 to 8–10 years of age and involves the ability to discern how an object, visible to another person, is perceived by them (visual) and to construct a spatial representation of what they perceive (spatial) [4,10,21]. Different tasks, such as three mountain task by Piaget [17] or turtle task [9], have shown that children younger than 4–5 years old were unable to engage in level 2. However, Moll and Meltzoff showed that 36-months-old's were significantly correct in responding to a level 2 test with color filters [12]. As a result, Moll et al. argue that the level of cognitive engagement affects children's performance in level 2 perspective taking tasks [13]. They differentiate between tasks that require confrontation and the ones that only require to take or adopt perspectives. Since children's performance is a function of task complexity, not just the perspective taking itself, we decided to run a pilot to discover the appropriate age for children to participate in our study. Our criteria included the ability to distinguish between left and right and being able to perform the essential task - giving instructions to the robot. However, we wanted children to be at a developmental stage where we can document their choice of perspective and evaluate their learning gain from the interaction.

A great deal of robots in education research has focused on evaluating them as learning companions [8], tutors [2,3,5] and learners [1,6,15] in educational settings. Assigning the robots to any of these roles is subject to the learning objectives and the robot's intelligence. These studies bring an understanding of how robots can be beneficial in educational settings, and the developments still needed. The main goal of this research is to approach the topic of robots in education by generating a decision-making model of perspective-taking for the robot inspired by children's behavior. To elaborate on both topics of perspective taking and robots in education, we have designed the following activity that simulates the collaborative interactions between the child and the robot with spatial perspective taking as a requisite to complete the task. To inform the future design of our perspective taking model and to ensure that we target the right age group, we ran a qualitative pilot study with 7 children from 4 different age groups. In this paper, we briefly describe the design of the task and interaction, our analyses of children's performance, the selection of appropriate age group, and what we learned from the pilot. As a result, we have formulated the following research questions for our pilot study:

RQ1: At which age group are children able to comprehend the task and carry it out without the help of the facilitator?

RQ2: At which age group are children able to correctly differentiate between their left/right and the robot's left/right?

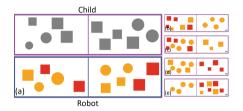**Fig. 1.** The experimental set-up with the child side activated (Color figure online)

Fig. 2. Medium level: (a) main task with (b) M1, (c) M2 (d) M3 (e) M4 goal cards (Color figure online)

2 Pilot Study

A total of 7 participants (4 female, 3 male) between the ages of 6 and 9 years old took part in this study. They were selected from four different age groups that were going to start 1st, 2nd, 3rd, and 4th grades. The study had received ethical approval from the university's ethics committee and parental consent forms were collected from the parents of the participants prior to the main experiment.

2.1 Study and Task Design

To design an activity involving perspective taking, we consider three concepts observed in the utterances with spatial perspective taking: frame of reference, perspective marking, and perspective taker's role. **Frame of reference** is a set of axes or origin points for addressing position of the objects or their spatial relationships [7,11,23]. Here, we mainly focus on *egocentric* (from the self point of view) and *addressee-centric* (from the other point of view). **Perspective marking** separates the utterances into *implicit* and *explicit* based on the existence of possessive adjectives in the sentence [19]. **Perspective taker's role** corresponds to the differentiation between the *speaker* or *instructor*'s and the *listener* or *manipulator*'s perspectives. Based on these concepts, if the robot tells the child "give me a brick on *your* right", the robot is addressee-centric, explicit, and an instructor/speaker. Children interacted with the robot in a short practice session (child as instructor) and 4 main sessions (child, robot, child, robot as instructor, respectively). We will be looking at children's understanding of the task, recognizing their own and the robot's left/right, and their overall performance.

For the task, we designed a simple game called the *objects game*, which includes moving circles and squares from one side of the screen to the other side, Fig. 1 shows the experimental setup with the child player. The main screen is composed of *squares* and *circles* in two colors: *red* and *yellow*. The game has two difficulty levels, which are a function of the color and shape of the objects presented in that level. Level 1 includes *yellow circles* and *yellow squares* (solved in 2 moves), while level 2 has the additional *red squares* (solved in 3 moves). The goal cards represent the desired final state of the game that players must recreate by moving the objects. Figure 2 shows the main game with 4 out of 6 available goal cards. When the game starts one player guides the other to reach the state represented in the goal card without directly showing it to them. The player with the goal card is called the *instructor*, and the player moving the objects is the *manipulator*. The instructions have three components: the color, the type of the object, and the moving direction. An example of a proper instruction is *"move the yellow circles to the right"*—an implicit utterance that can be either egocentric or addressee-centric.

3   Discussion and Conclusion

To determine the appropriate age group for participating in the main study, we look at two criteria: children's ability to understand the task and to differentiate between their left/right and the robot's left/right. We want children to be able to understand the central concept, be challenged by the difference in perspectives, and make a decision to deal with the difference, either successfully or not. During the interaction, we noticed two participants (6 and 7 years old) had fundamental problems distinguishing between their left/right. Furthermore, the 6 years old child had problems identifying the shapes to produce the instructions. We had a plan to accommodate children with left/right issues by putting stickers on their hand and the robot's hand. Several psychology studies have used this technique in their perspective taking studies [16]. However, it did not solve those children's issues and they were still confused about the robot's difference in perspective. We discussed this issue with the teachers, who advised us that the task was too difficult for children starting 1st and 2nd grades. On the other hand, we observed acceptable performances from children in 3rd and 4th grade. The children in the 3rd grade were able to comprehend the task, they were egocentric at first, but one of them managed to recognize the discrepancy between theirs and the robot's perspective and update their instructions. With 4th grade children, we observed that they effortlessly recognized the robot's different perspective and update theirs. Based on our observation of children's performance and further discussions with the teachers, we decided to select children at 3rd and 4th grade. We excluded younger children due to their issues with left/right and understanding of the task.

Furthermore, we were able to recognize a shortcoming in our interaction that was affecting children's perception of the robot. During the interaction, when the child instructed the robot in implicit egocentric instructions, considering the robot's egocentric perspective, the outcome of the move was opposite of

the child's expectation. In such cases, some children were expecting the experimenter to explain why, and most just assumed the robot was faulty. To prevent this, we decided to add some level of transparency to the interaction for the future experiment by making the robot ask for feedback after every move, and in response to a negative feedback convey its egocentric perspective (e.g. "but I moved them to my left/right"). Using the takeaways from the pilot in our next study we plan to explore *how children's choice of perspective is affected by the robotic's choice of perspective.*

References

1. Yadollahi, E., Johal, W., Paiva, A., Dillenbourg, P.: Anonymous: when deictic gestures in a robot can harm child-robot collaboration. In: Proceedings of the 17th ACM Conference on Interaction Design and Children, pp. 195–206 (2018)
2. Belpaeme, T., et al.: L2tor-second language tutoring using social robots. In: Proceedings of the ICSR 2015 WONDER Workshop (2015)
3. Castellano, G., et al.: Towards empathic virtual and robotic tutors. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS (LNAI), vol. 7926, pp. 733–736. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39112-5_100
4. Flavell, J.H., Everett, B.A., Croft, K., Flavell, E.R.: Young children's knowledge about visual perception: further evidence for the level 1-level 2 distinction. Dev. Psychol. **17**(1), 99 (1981)
5. Gordon, G., Breazeal, C.: Bayesian active learning-based robot tutor for children's word-reading skills. In: Twenty-Ninth AAAI Conference on Artificial Intelligence (2015)
6. Hood, D., Lemaignan, S., Dillenbourg, P.: When children teach a robot to write: an autonomous teachable humanoid which uses simulated handwriting. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, pp. 83–90 (2015)
7. Levinson, S.C.: Language and space. Ann. Rev. Anthropol. **25**(1), 353–382 (1996)
8. Lu, Y., Chen, C., Chen, P., Chen, X., Zhuang, Z.: Smart learning partner: an interactive robot for education. In: Penstein Rosé, C., et al. (eds.) AIED 2018. LNCS (LNAI), vol. 10948, pp. 447–451. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93846-2_84
9. Masangkay, Z.S., McCluskey, K.A., McIntyre, C.W., Sims-Knight, J., Vaughn, B.E., Flavell, J.H.: The early development of inferences about the visual percepts of others. Child Dev. **45**, 357–366 (1974)
10. Michelon, P., Zacks, J.M.: Two kinds of visual perspective taking. Percept. Psychophys. **68**(2), 327–337 (2006)
11. Mintz, F.E., Trafton, J.G., Marsh, E., Perzanowski, D.: Choosing frames of reference: perspective-taking in a 2D and 3D navigational task. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 48, pp. 1933–1937. SAGE Publications Sage, Los Angeles (2004)

12. Moll, H., Meltzoff, A.N.: How does it look? Level 2 perspective-taking at 36 months of age. Child Dev. **82**(2), 661–673 (2011)
13. Moll, H., Meltzoff, A.N., Merzsch, K., Tomasello, M.: Taking versus confronting visual perspectives in preschool children. Dev. Psychol. **49**(4), 646 (2013)
14. Moll, H., Tomasello, M.: Level 1 perspective-taking at 24 months of age. Br. J. Dev. Psychol. **24**(3), 603–613 (2006)
15. Muldner, K., Lozano, C., Girotto, V., Burleson, W., Walker, E.: Designing a tangible learning environment with a teachable agent. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS (LNAI), vol. 7926, pp. 299–308. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39112-5_31
16. Newcombe, N., Huttenlocher, J.: Children's early ability to solve perspective-taking problems. Dev. Psychol. **28**(4), 635 (1992)
17. Piaget, J.: Child's Conception of Space: Selected Works, vol. 4. Routledge, Abingdon (2013)
18. Sodian, B., Thoermer, C., Metz, U.: Now I see it but you don't: 14-month-olds can represent another person's visual perspective. Dev. Sci. **10**(2), 199–204 (2007)
19. Steels, L., Loetzsch, M.: Perspective alignment in spatial language. arXiv preprint cs/0605012 (2006)
20. Surtees, A., Apperly, I., Samson, D.: Similarities and differences in visual and spatial perspective-taking processes. Cognition **129**(2), 426–438 (2013)
21. Surtees, A., Samson, D., Apperly, I.: Unintentional perspective-taking calculates whether something is seen, but not how it is seen. Cognition **148**, 97–105 (2016)
22. Tomasello, M.: Origins of Human Communication. MIT Press, Cambridge (2010)
23. Trafton, J.G., Cassimatis, N.L., Bugajska, M.D., Brock, D.P., Mintz, F.E., Schultz, A.C.: Enabling effective human-robot interaction using perspective-taking in robots. IEEE Trans. Syst. Man Cybern.-Part A: Syst. Hum. **35**(4), 460–470 (2005)