Interaction Design and Architecture(s) Journal - IxD&A, N.44, 2020, pp. 109 - 128

Demystifying Robots in the Co-Design of a Tutee Robot with Primary School Children

Wolmet Barendregt[1,3], Sara Ekström[2], Svea Kiesewetter[1], Lena Pareto[2]
Sofia Serholt[2,1]

[1] Department of Applied IT, University of Gothenburg, Forskningsgången 6, SE-417 56, Gothenburg, Sweden
[2] School of Business, Economics and IT, University West, Gustav Melins gata 2, SE-461 86, Trollhättan, Sweden
[3] Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, De Zaale, 5600 MB, Eindhoven, The Netherlands
w.barendregt@tue.nl

Abstract. This paper contributes to the existing body of knowledge on the co-design of novel technologies with children. As part of a three year research project aiming to design and develop a robot tutee for use in mathematics education, we present the initial phases of our design approach with children, in which we draw on principles of Participatory Design and Co-design. As part of the early stages of this process, we included a demystifying phase (I), and a gradual introduction to the robot's capabilities (II), in order to foster reasonable expectations in children and gather feasible design input. Drawing on the Time-Space-Structure framework, two primary schools were involved in the co-design process, where children in grades 2 and 4 participated in a set of workshops. We discuss the benefits and tensions of our approach, and reflect on its implications for mutual learning, hoping to inspire further exploration in this field.

Keywords: co-design, robot tutee, children.

1 Introduction

Co-design, or Participatory Design, has a long tradition rooted in the Scandinavian social and political movements related to the transformation of workplaces during the 1970s. From around the end of the 20th century it has also become quite common to consider children as participants in participatory design processes, especially through Alison Druin's extensive emancipatory work in this area [1-3]. Since this time, children have been involved successfully in, e.g., the design of exhibits [4] and educational games [5].

Currently, robots, and social robots in particular, are entering all aspects of society, such as hospitals, airports and education. However, the "responsibility for making decisions about the design of such robot applications has so far been largely in the hands of robot designers" [6]. Based on their team's experiences of including, e.g., elderly and patients [7] and office workers [8] in robot design activities, Lee et al. [6] propose to use a more inclusive, participatory design process for the design of social robots. As recent research suggests, attitudes may become more positive and anxiety towards educational robots may reduce by involving adult students in prototyping processes of (hypothetical) educational robots [9]. However, involving children in the co-design of an educational robot is still not very common. In this paper, we therefore aim to describe our experiences of involving children in the design of a robot tutee. By teaching the robot tutee how to play an arithmetic game, and by answering the inquisitive questions of the robot tutee [10], children are supposed to improve their arithmetic comprehension through mechanisms of learning-by-teaching [11]. While the end result of this design process (the fully developed robot tutee) may be interesting as well, here we focus mainly on children's involvement in the design process of this robot tutee.

We will start this paper with an overview of the literature on co-design of (educational) robots with children. Thereafter, we will present our particular context where we aim to design a robot tutee to play an arithmetic game with children, as well as our co-design approach, which pays particular attention to children's learning opportunities, especially related to how robots works, that are hallmarks of participatory design processes [12]. The core of this paper is a description of our experiences during the first phases of our project, and especially children's evolving understanding of, and attitude towards, a classroom robot during this process. We will also point out some of the difficulties and advantages of our approach, hoping to inspire others in their particular endeavors to include children in the co-design of educational robots.

2 Background

2.1 Co-design of (Educational) Robots with Children

One of the first projects involving children in the design of a robot, was performed by Druin and her team [13]. In this project, researchers and children between 7 and 11 years old worked in intergenerational design teams to develop a storytelling robot. The design teams started by trying out robots in a robotics lab, and asking questions of the researchers in the lab. They all took notes about the experience and upon return to the researchers' HCI lab, they wrote down what they liked and did not like about the robots they had seen. Thereafter, three intergenerational teams worked on building prototypes for the skeleton, the skin and sensors, and the software for a storytelling-robot. Based on their experiences in this project, the researchers provided three general guiding principles for co-design with children.

The first principle is that there is a need to develop new power structures between adults and children. They specifically established this change in power structures by not allowing anyone to raise their hands, and by giving both children and adults an active role in discussions. The second principle is that all design partners must have a voice in the design process. While they admitted that this was not an easy feat, the researchers found that drawing, writing, and building could be used to capture ideas, and the artifacts produced during these activities could serve as 'a catalyst and bridge for discussion'. The last guideline was the need to create a comfortable design environment. In their project, they established this by not allowing anyone to dress formally, having everyone sit on the floor or in beanbag chairs, and using low-tech materials. While these findings have been used and refined in numerous co-design projects with children since then, they are not very specific for the design of (educational) robots with children.

Obaid and colleagues have worked specifically on the involvement of children with and without robotic knowledge in the early phases of a robot design process. They started with an exploratory study to determine what different design elements children consider concerning the design of classroom robots compared to those of adult interaction designers [14]. In their study, they tasked small groups of either children or interaction design students with designing a classroom robot. An analysis of the drawings and the discussions during the design process revealed some differences in robot designs between children and adult designers. While interaction designers envisioned a small or child-sized non-gendered animal- or cartoon-like robot, with clear facial features to express emotions and social cues, children envisioned a bigger human-machine robot. Based on this first study, they developed a design toolbox (Robo2Box) [15] that included the most common design elements found in both the drawings and the focus group discussions (from both children and interaction designers), in addition to several design elements identified by Woods [16]. Their intention with the Robo2Box was to support children's involvement in the co-design of a classroom robot by removing the need to rely on drawing skills, and by offering children inspiration from other children, the literature, and interaction designers. While this toolbox could be useful to help children imagine some of the form factors of their classroom robot, it is not meant to address other characteristics of an educational robot, such as its behavior. Furthermore, they never used the toolbox in an authentic robot design project with all of its practical limitations. While their work is certainly relevant for the co-design of classroom robots in general, for our purpose, the use of the Robo2Box was not that relevant as the robot embodiment was decided upon before the co-design process started. This constraint will be explained in more detail in Section 2.3.

Arnold, Yip, and Kung [17] performed and described one co-design session with children to develop a friend robot. In this session, small groups of children between 6 and 11 years old used a well-known co-design design technique called Bags of Stuff where they received a bag of art supplies to create a low-tech prototype of a friend robot. During the session, the adult researchers assisted the children and gave suggestions while building their own robots. Based on this design session the researchers were able to pinpoint four important aspects of friend robots for children. However, similar to the work by Obaid et al. [10], their intention was not to design an actual robot with which the children would be able to interact. Rather, they were interested in understanding what kind of design input children could give.

Finally, Bertel et al. [18] developed a participatory design framework for implementing educational robots in real-world learning environments. This framework includes Time, Space, and Structure (TSS) as important factors for planning participatory design activities to develop educational robots in schools. They claim that the balance between these three elements determines the level of user participation and the nature of the research questions that a particular user study can address. *Time* relates to the length of the study in question. They argue that while shorter studies can be used to study users' first impressions and initial attitudes towards educational service robots, longer-term studies are necessary to do 'research and development iterations as well as ongoing adjustment of the project according to newly gained knowledge and experiences and thus more reliable and replicable results' (p. 439). *Space* relates to several different aspects, such as where the study is conducted, but also how far developed the robot prototype is, and the creation of a safe and creative space where participants can share their personal views and experiences. Finally, *structure* covers the level of control in the study, e,g., whether it is a controlled experiment aiming for validity and generalizable results, or a creative process. They argue that 'too much control can demotivate the participants or even induce reluctance towards the researchers or the technology. On the other hand a total lack of control could possibly leave the development unfocused, participants insecure and unmotivated, and the technology unused' (p. 440). They applied this framework in a case study on the use of the therapeutic robot seal Paro at a school for children with autism. Their aim with this study was to 'explore the potential of robots to promote motivation, communication, play and learning in autism education, to investigate the importance of structure' [in the proposed framework, and] 'to identify additional contextual factors that both influence and are influenced by the technological intervention' (p. 440). However, while the framework was developed for participatory design practices in real-world learning environments, the case study focused mainly on teachers' experiences of the co-design process. Although the design teams involved both teachers and children, the teachers mainly observed the children when interacting with Paro in order to develop different didactic designs to be shared with and used by others.

Altogether, we argue that there is only limited work related to projects where the designers or researchers aim to co-design a real educational robot with children, taking into consideration all the practical limitations and constraints of such a project. Indeed, given those limitations and constraints, and the need to work with the children over a long period of time, we are currently only able to provide the reader with the insights from the first year of our co-design project, admitting that there is still a long way to go before our robot will find its way into a classroom. However, we do see the need to share our experiences given the limited number of studies addressing co-design of an educational robot as part of a real design project.

2.2 Co-Designing a Robot Tutee for an Arithmetic Game

In our project, we aim to design and implement the behavior of a Pepper[1] robot, which will act as a tutee together with a validated mathematics game including a teachable agent. The idea builds on a promising approach using robots as teachable companions in a learning-by-teaching paradigm, where the student becomes engaged in the act of teaching someone else. Research has shown that a learning-by-teaching approach with a teachable agent may have some educational benefits, e.g., [19, 20]. For example, Chin et al. [21] showed that teachable agents not only help students learn the material that they teach the agent, but also prepare them to learn new content when the technology is no longer used. According to Blair et al. [20], teachable agents provide opportunities

---

[1] A humanoid robot created by Softbank Robotics

to optimize learning-by-teaching interactions by adopting the following four principles: to use explicit and well-structured visual representations; to enable the tutee to take independent actions; to model productive learner behaviors; and to design learning environments that support teaching interactions. Designing for learning-by-teaching thus needs to focus on how the teaching activity can optimize the benefits of teaching for the human tutor as opposed to approaches that focus on how the artificial tutee can learn, i.e., machine learning aspects.

With a few exceptions, most teachable companion systems use virtual agents rather than physical robots [22]. However, there may be several advantages of using robots rather than virtual agents, such as the physical presence strengthening the perception of having a social partner and being more engaging to students. Some research has also shown that students may learn more from embodied interactions [22]. The perceptions of machines, in particular robots, are influenced by anthropomorphism [23], and humans tend to apply social rules to machine-based companions even when they know they are dealing with machines [24]. However, designing robotic companions that people find socially appealing and engaging is a great challenge, and aiming for autonomous behavior means continually taking actions that are reasonable and useful as well as believable and engaging in each particular situation [25]. For example, according to Zhang and Sharkey [26] perceived emotional expressions on a robot face can be affected by situational parameters such as the context. Dautenhahn [23] therefore proposes that robot designers need to focus on the social aspects of interaction as much as task-related issues in order to make robots more believable in behavior. However, following Lee's [6] advice, we argue that this should entail performing co-design activities with the children.

2.3 Constraints of the Project

Although the project aims to develop a functioning robot tutee for use in the primary school classroom, there were several constraints that influenced our co-design process. We will discuss these constraints below.

2.3.1 No Focus on Embodiment of the Robot

Firstly, unlike in some of the co-design studies described above, we did not have the opportunity to determine the embodiment of our robot as part of the co-design activities with the children. Given the research funders' requirement to apply for funding of specific technical equipment and the university's need to purchase any equipment as soon as possible after the project start due to depreciation issues, we had to make a decision about the embodiment before being able to involve any schools. Having some previous experience with a Nao robot, we decided to purchase the more advanced Pepper robot, which still uses a similar platform. The Pepper robot also has the advantage that it is slightly bigger than a Nao robot, which fit better for the purpose of our project, as will be described below.

2.3.2 Existing mathematics game as point of departure

Second, our project builds further upon a graphical arithmetic game developed in a previous project by one of the authors. In this game, children can teach a virtual teachable agent to play the game on an interactive whiteboard. The original mathematics game teaches arithmetic, specifically base-ten concepts [27]. The game offers a set of two-player mini-games with basic content (addition up to 10) to more advanced content (multiplication and division using positive and negative numbers up to 1000). The game is a combined card- and board-game, and the players take turns choosing cards for the common game board. Each player has a card-hand representing numbers, and the player chooses an arithmetic operation. This way, a game sequence represents a sequence of arithmetic computations. The mini-games have different goals that reward points to the players, e.g., to combine cards that add to 10, or to maximize the number of carry-overs in a computation. Players can choose to collaborate or compete. Instead of using traditional mathematical representations as numbers and symbols, the game uses graphical and animated representations of numbers and arithmetic operations: integers are colored blocks; arithmetic operations are animated actions involving those blocks. For example, ten's are orange-colored boxes containing exactly ten red blocks, and the blocks are explicitly packed and unpacked in the boxes illustrating carry-over operations. The graphical model ensures all mathematical rules, which allows the player to discover and explore properties and relations since possible actions are mathematically valid. In the original game, children can teach a virtual tutee to play the game. The tutee starts with no knowledge meaning it will choose cards by random if played alone, i.e. it will "guess". After being taught, the tutee will mirror the child's playing behavior that is automatically analyzed in the game. This means that the tutee will choose cards at a similar computational and strategic knowledge level as the child played when teaching, and the tutee will only progress into more difficult questions when the child has answered the basic questions in a convincing way. Thus, the tutee behaves as if it learns from its tutor and it adheres to the principles of a teachable agent in [20]. The child as tutor teaches by playing the game while the tutee "observes" the tutor's actions or suggests own actions and then poses written, multiple-choice questions related to the tutor's just-performed actions. The purpose of the tutee questions is to stimulate the child to reflect on her own actions and mathematical thinking, but still be questions that an inquisitive playing partner could ask. Controlled classroom studies of the game augmented with this virtual tutee have shown significantly higher learning gains for students playing the game compared to non-playing controls concerning base-10 concepts. The studies have also shown that children in early education can become engaged in advanced mathematical thinking, and are able to act successfully as tutors [28]. Although the arithmetic game is an integral part of our setup, we do not primarily aim to redesign the game with the children. However, we are of course open to suggestions from the children.

The aim of our present project is thus mainly to co-design both the social and task-related behavior of the Pepper robot that can replace the virtual tutee in the original game. A picture of the envisioned setup is shown in Figure 1.

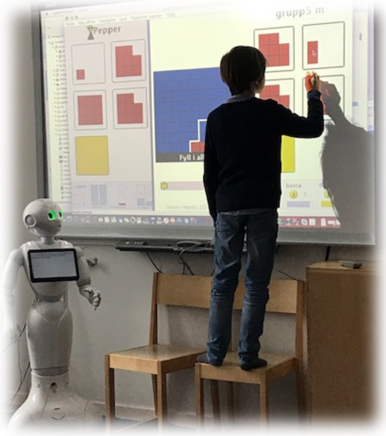

Fig. 1. The envisioned setup of our robot tutee playing together with a child on the classroom interactive whiteboard.

Given all those parameters, we set out to involve schools in our co-design process, which we will describe in more detail in the subsequent section.

3 Our Co-design Approach

3.1 Understanding and Demystifying Robots

Specifically related to the participatory design of AI systems, such as robots, Bratteteig and Verne [29] have pointed out that non-expert co-designers need to have a basic understanding of what such systems can and cannot do. In the case of children co-designing an educational robot, we felt that this was particularly necessary since they may have very high, but unrealistic, expectations of robots based on robot portrayals in movies and books. Furthermore, we reasoned that in order to balance the power relations, as suggested by Druin [13], the children also needed to have an understanding more similar to that of the researchers of how robots function. Finally, we assumed that there would occur a particular kind of novelty effect where children would initially be very enthusiastic to participate in the design process because of the novelty of the robot.

We therefore explicitly dedicated a part of our design work towards teaching the children what robots are, and how they are developed. Ultimately, we aimed to create co-design activities for mutual learning in which children would be able to gain knowledge about robots and the practice of designing a robot's behavior, and where we as researchers, at the same time, would gain useful knowledge about children's preferences and ideas for a robot tutee. This dual aim created a need to deal with a delicate balance between teaching the children about robots, while not hampering their imagination of what they would like a robot tutee to do.

To this end, we set up two series of design workshops in the schools (each workshop was held with children in all four classes). Although we intend to hold several more workshops with the same children throughout our project, these two series of workshops constitute our progress for the first year of the project, focusing specifically on demystifying the robot and introducing children to the idea of a robot tutee.

Each workshop was held with all the children that had received written consent from their legal guardians allowing them to participate. They also signed a child-friendly consent form themselves at the start of each workshop. Workshops could usually not take more than around one hour per group, as the teachers had to plan them as part of the regular school day. To participate in the workshop, children left their class in small groups to attend each workshop, meaning that there were no teachers present during the workshops. Each workshop was run by three researchers. Working with small groups allowed all children to have the possibility to have a say, as advocated by Druin. By holding the workshops without the teachers present, we hoped to get away from the existing power structures within the school context.

3.2 The Time-Space-Structure Framework

In terms of Bertel et al. [18]'s TSS (Time-Space-Structure) framework mentioned earlier, we aimed to involve children and teachers for the whole duration of our project, meaning four years. We therefore started by contacting some of the schools that had shown an initial interest in our project, due either to their experience with the arithmetic game or due to promotion of the project at the university website. After initial workshops in several schools and with several teachers, two schools were chosen to continue to work with us based on an interest and initiative shown. School A contained two second grade classes, and School B contained two fourth grade classes. During our work with the children, we did not collect any demographic data, but the typical age for second grade in Sweden is 7-8 years old, and for fourth grade it is 10-11 years old.

Related to the space dimension of the framework, there are several advantages to working in the school setting as opposed to the lab, such as the familiarity of the surroundings to the children and no need for costly and time-consuming travel. Therefore we aimed to perform our co-design activities in the school setting. However, when working in the classroom, we also had to deal with the fact children may not consider the school environment a creative space, and that they may easily fall into their role as students rather than active co-creators, especially when teachers are around [30]. We therefore aimed to work with the children (and teachers) separately. The need to move our Pepper robot to the schools was a slight disadvantage due to its large and cumbersome size, while it also obliged us to make use of the schools' sometimes unreliable technical infrastructure. However, we felt that the advantage of working in a familiar environment and not requiring the children and teachers to travel outweighed this disadvantage. Furthermore, the classroom is the authentic setting in which the robot-enhanced educational game is envisioned to be used.

Related to structure, we did not aim to do any controlled experiments. Instead, we considered children to be experts on their own lives in school, with knowledge about how they would like a robot tutee to behave. We thus wanted the children to have the freedom to share this knowledge with us, but without making them insecure about what to do. We were also aware of the need to involve the children in a way that would fit the school structure, and that would be acceptable to the teachers as a school activity. As argued for by several researchers, e.g. [12, 31, 32] considering the gains of participation for all parties involved is an important aspect of doing co-design.

3.3 Activities Performed

Table 1 presents the different activities undertaken as part of our co-design approach. In this table, we can distinguish two phases when working with the children, **Phase I: Demystifying robots** and **Phase II: Introducing children to the robot as tutee**. During the first phase we mainly aimed to engage the children in activities to teach them about how robots function. We also hoped that this first phase would help to temper some of the novelty effect. After the first phase, we implemented Pepper with some social behavior based on the output from the first phase and connected it to the arithmetic game using the dialogue mechanism from the virtual tutee, in order to provide a more tangible and experiential illustration of its capability. In the second phase we aimed to introduce the children to Pepper as a robot tutee that can play the game. By gradually introducing the children to our setup, we aimed to prepare them for further co-design work to develop the robot tutee's behavior.

Table 1 Activities undertaken as part of our co-design approach

Phase	Activity
Phase 1: Demystifying robots	1a Writing post-it notes on thoughts about robots
	1b Group discussion about robots
	1c Introduction to Pepper
	1d Creating dialogues for Pepper
	1e Semi-structured group reflection on workshop
	1f Post-workshop questionnaire
Implementing the robot as a tutee connected to the arithmetic game	
Phase 2: Introducing children to the robot as tutee	2a Interacting with Pepper as robot tutee in the game
	2b Semi-structured group reflection on workshop
	2c Post-workshop questionnaire

Following the advice of Barendregt et al. [33] to increase the user gains by allowing time for reflection, both workshop series ended with a short group reflection where we asked the children what was unexpected or new to them and what they had learned. We also sat down as a team after each workshop in the first series to discuss our own reflections about what we had learned, both in terms of the robot design and in terms of the workshop design. Unfortunately, we did not have time to do such a reflection round as a team for the second phase due to the tight schedule in the schools. The reflections about the second workshop phase described in this paper are therefore mainly based on later testimonies of the activity, as well as the data collected during the workshops. Roughly one or two weeks after each workshop, the children received a post-workshop questionnaire. The questions in this questionnaire were based on our experiences during the workshop and aimed to catch children's further reflections and suggestions after a short period of rest. As argued by Serholt [34], conducting follow-up studies with children on their experiences of interacting with robots after some time can be way to circumvent the novelty effect of the intervention while at the same time allowing for research instruments to be tailored specifically to what occurred during the actual interactions or study.

3.4 Research Methods and Data Analysis

We applied a mixed-methods approach for this research, gathering written post-it notes from the children, audio recordings from all design sessions and group interviews, as well as quantitative data from the post-workshop questionnaires. We also took observation notes during all interactions and audio-recorded the post-workshop reflection session between the researchers for the first phase. All recordings were transcribed and one of the authors then used all available qualitative materials to create themes. These themes were then discussed with the other authors to determine whether they represented their impressions of the interactions with the children during the workshops. Finally, we connected the qualitative findings to the quantitative results from the questionnaires to be able to provide a complete picture of our design process.

4 Our Experiences

In this section, we will describe the two phases in more detail, integrating some of the outcomes of the workshops with reflections on how well the different activities served our purpose of mutual learning. We do not discuss all outcomes in detail, since that is beyond the scope of this paper.

4.1 Phase 1: Demystifying Robots

In the first phase, involving 73 children, 30 in second grade and 43 children in fourth grade, we mainly aimed to demystify robots by teaching children about how they function. We also hoped that this first phase would allow the children to get used to the robot so some of the novelty effect would wear off.

Our first activity with the children was a post-it note activity where the children individually wrote down their thoughts when hearing the word 'robot'. As expected, related to the novelty effect, many children wrote down words that expressed their feelings of enthusiasm about robots or about going to meet a robot, such as "cool", "fun", "exciting", or: "Cool! I have never met a robot before and now I will". The post-it note activity also showed that the children either addressed robots as deviating from a human, such as it being a "fake human" (Second-grader), or a "human but made of metal" (Fourth-grader), while other mainly addressed robots as a machine with human qualities, such as "a machine that is alive" (Fourth-grader). A few children showed an understanding of the connection between programming and robots, as one child wrote, that a robot can be "programmed and developed" (Second-grader) while another wrote that a robot is "a machine that can be programmed for a thing that it can do" (Fourth-grader). The children also wrote down some expectations about what a robot would be able to do. Not surprisingly, their expectations of these capabilities where rather high, such as the robot being smart, knowing a lot, and being friendly and nice. Even though this activity was meant to be done individually, it was observed that children listened to each other and in some cases copied answers from their classmates. In some cases, the children had already seen Pepper when one of the other groups had done the workshop, which influenced their expectations—these encounters were brief and accidental. As we chose to conduct our study in actual schools, we had to deal with the fact that schools are an uncontrolled environment, at least from a scientific perspective. Indeed, researchers need to consider fairness aspects in the sense that all children who have the desire to participate should be allowed to do so. Typically, one also needs to consider the value for teachers around organization, and that it might be difficult in practice to keep some children away from an ongoing (exciting) intervention happening close-by.

The post-it note activity, where the children worked individually, was followed by a group discussion where the children could talk about what robots they already knew and discuss further about what robots are and what they can or cannot do. As we had expected, most children had seen robots in popular media, such as films and games, but some children also mentioned that they had read about robots in books at school, and one child had even programmed a robot mouse in a pre-school class. During the group discussion, the children and researchers engaged in a conversation about what robots are. One way to discuss this was by using the examples of robots they had encountered before. Based on their previous experience with robots, several children were aware that robots need to be programmed, both for its physical and cognitive abilities. However, many of them defined robots in terms of physical attributes resembling humans, such as having legs, eyes, a head, or fingers.

Even though most children engaged actively in the group discussion, they were clearly aware that they were going to meet a robot soon, which sometimes made them less focused. Indeed, one of the children could not wait any longer and asked during the discussion: "Can we meet the robot now? Please, please, please" (Fourth-grader). This part of the workshop became actually harder and harder to perform during the day, as children from previous groups informed the children that had not participated yet, about what was going to happen. This relates to the tension mentioned previously regarding the post-it note activity, yet, here, the issue could have been addressed by conducting parallel sessions to avoid children influencing each other. In most research projects on educational robots, however, it is not feasible to acquire a set of robots for obvious reasons such as their high costs and need for maintenance.

When it was time to meet Pepper, we intended to use this time to, e.g., show the children how the robot's sensors work. The robot was therefore programmed with a few dialogues where Pepper asked the children about their school and favorite school subjects, using the built-in options to let it, e.g., track people with its gaze and understand very basic words. One pre-programmed dialogue included a piece of saxophone music to which Pepper danced.

There was a large variation in how the children approached the robot. A few children did not want to go close to the robot at all and made sure to keep a distance to it. They also admitted they thought Pepper was a bit scary. Other children were extremely physical with the robot and kept hugging and touching it and did not want to leave. However, they seemed aware that the robot is fragile so they were rather gentle with it. This was probably influenced by the relatively short-term nature of the direct physical encounters between the participants and the robot, as well as the presence of the researchers and possible instructions of the teachers. Other children focused on the interaction trying hard to get Pepper to listen and respond to them. In some cases, the eagerness of some of the children obstructed the less assertive children from interacting with it. In some cases we thus had to actively try to give the less assertive children the possibility to interact by engaging the other children in some other activities. Furthermore, our intention to really show how a robot works was in some cases overshadowed by children's enthusiasm to get Pepper to respond. However, there were also several groups that really wanted to understand how Pepper technically functions, e.g., how it can see or hear anything. In many cases, it was clear that the interaction with Pepper did help the groups get a feeling for what it entails to make a robot do whatever it does, as exemplified in the following transcript:

- Have you bought this one? How many millions?
- 'It'
- But it looks like a 'her'
- What happens if it topples over?
- Does it feel anything?
- Does it hurt if it can shake hands? Does it squeeze?
- Can it talk?
- It looks like it is green. Is it listening then? Then it stands a little bit like…
- It likes to follow you. It likes to look at you
- Do you have to talk this clearly?
- Now I know, they program to click. They put something in, in this way... something it should do. You try to program… So you can put in, what it should do
- O, that's a lot of work
- So hard
- I don't get a thing

After this first introduction to Pepper, it was time for the children to create their own dialogues for Pepper. In order to do so, each group received two sheets of paper, one for the introduction, and one for farewell. Each sheet contained two columns, one for Pepper and one for the child. After filling out a sheet, the groups could ask one of the researchers to help them implement their dialogue on Pepper and test it. We were pleasantly surprised that most groups were very good at writing those dialogues. Implementing the dialogues on Pepper required some time. However, since the children were able to see exactly how the programming was done on the whiteboard connected to the computer, the children usually understood what was going on and had the patience to wait for the end result. Figure 2 shows one of the groups creating a dialogue for Pepper.

Fig. 2. A group creating a dialogue for Pepper.

Once a dialogue was implemented, the children could test it in interaction. Many of them quickly picked up that they had to speak clearly to Pepper to make themselves understood, and that they had to say the exact words they had used in their dialogue to allow Pepper to recognize them correctly, so they used their written dialogue as manuscript to read from. However, some children did not want to talk to Pepper at all, and some children gave up on talking to Pepper when it did not understand what they were saying the first time. Pepper displayed problems understanding some children even though they said the correct word loud and clearly. Two groups seemed more vulnerable to this problem; the children with a different mother tongue than the one used in the school, and girls (with either Swedish or another mother tongue) more than boys. However the robot behavior was not consistent enough to draw any conclusions yet, and we intend to investigate the matter further by analyzing the children's voices in relation to the robot behavior from the video-recordings. We also discussed to lower the threshold for when Pepper determines whether it recognizes certain words in the second phase. Regardless, the bias of existing speech recognition software is clearly an issue when considering the multiculturalism of society. That some children would be excluded from interacting with an educational robot due to their accents is extremely problematic and risks impeding an inclusive educational environment. This clearly needs to be recognized by the research field and/or the companies developing social robots for verbal interaction.

Towards the end of the workshop, one of the researchers sat down with each group to reflect on their experiences during the workshop. We were especially interested in what they had learned about robots and whether we had succeeded in demystifying robots in some way. Some groups were very good at expressing what they had learned, while other groups struggled with talking about their experiences. Generally, two types of reactions could be discerned, which were both related to the fact that Pepper was very minimally programmed. One typical reaction was to be slightly disappointed about Pepper's limited capabilities. The children had wanted it to be able to do all kinds of dances and play music. The other reaction, however, was that it was fun that Pepper was not able to do so much, because this made it possible to decide what Pepper should be able to do and teach it. In both cases, we felt that this understanding served our purpose of helping the children to understand what robots are, and slightly dissolve the mystery around them.

At the end of the first phase, we wanted to hear the children's reflections about their experiences during the workshops. Furthermore, as a preparation for the next phase, we wanted to get some insight into their thoughts about learning companions, and robotic learning companions in particular. We therefore sent out a questionnaire to all children about a week after the workshop asking them what they thought a robot was, what they thought about Pepper, how they would like a human learning companion to behave and not behave, and whether they would like a robot tutee to behave similar to a human learning companion or not. For this post-workshop questionnaire using open-ended questions, we received answers from 58 children. These were analyzed qualitatively by searching for themes within the answers [35]. The detailed results of this analysis are provided in another paper [36]. However, one of the most interesting results was that several children indicated that a robot tutee should behave as a better version of a human learning companion, meaning that the robot should be kind to everyone, give more compliments, and explain in a better way than a human. Many children showed a clearly better understanding of what a robot is, mentioning that is has to be programmed to do what it does. Finally, while many children wrote that they thought that Pepper was friendly and fun, some children repeated that they were slightly disappointed with its limited abilities. However, as one of the children wrote: 'It was good, but I know that it is not ready, and I believe that it will be good once it is finished'.

4.2 Connecting Pepper to the Math Game

After the first phase, we worked intensively on implementing a version of Pepper that could interact with the original mathematics game so that children could experience the robot in its intended role as a tutee. We used the original dialogues from the game's teachable agent as a point of departure, but we complemented this with some social behavior, as obtained from the first workshop. We also made sure to implement many different ways of answering the questions so Pepper would be able to recognize children's answers. An example of a dialogue that the robot could now have with a child is presented in Figure 3.

Check the child's understanding by guiding a reasoning sequence:

Robot: Will the card that you chose with 5 blocks give us a point?

- **Child**: Yes → **Robot**: So, the card 5 together with the 6 on the game board will be 10?
  - **Child**: Yes → **Robot**: Okay, I did not think so
  - **Child**: No → **Robot**: But doesn't it have to become 10 for us to get points?
- **Child**: No → **Robot**: So the card 5 together with the 6 on the game board is not 10?
  - **Child**: No → **Robot**: Good, then I know
  - **Child**: Yes → **Robot**: Okay, you know best
  - **Child**: Don't know → **Robot**: I don't know either, let's play!

When the robot knows more, it gives own suggestions and questions the child's choice:

Robot: I thought I'd choose the card with 3 blocks, why is 4 better?

- **Child**: (10 together, OR will give points) → **Robot**: So the 4 card is a 10-buddy with 7?
  - **Child**: No → **Robot**: But isn't the 3 card a 10-buddy with 7?
    - **Child**: Sure → **Robot**: So my card is better, right?
    - **Child**: Yes → Robot: Thanks buddy

Prompt ideas:

Robot: Now I think I get it! So you have to choose a card that will add to exactly 10 with the red blocks on the game board?

Fig. 3. Example of a different dialogue paths implemented in Pepper.

4.3 Phase 2: Introducing the Children to the Robot as Tutee

In total, 80 children participated in the second phase, 35 children in second grade, and 45 children in fourth grade. The main aims of the second phase were to introduce the children to the robot behaving as a tutee in connection to the arithmetic game and to determine it was possible for the children to engage in a dialogue around the game.

Once more, we worked with small groups to allow all children to engage with the robot and the game. Each group entered the workshop space where Pepper was prepared to play the game using the interactive whiteboard. As anticipated, the atmosphere during this second phase was slightly calmer than during the first phase, probably because the children had encountered the robot before.

In total, the children played 702 turns in the game, meaning that on average they chose and played 8-9 cards each. About every second turn, Pepper asked a question about the game or about their decision to choose a particular card. In the game logic there is a correct answer to each question, but this is only used to decide which questions to ask next, it is not revealed to the tutor and Pepper's responses are designed to be reasonable either way the child answers the question. As revealed in the game logs, Pepper asked 326 questions, and the children answered 218 of these answers correctly and 39 wrongly. However, the children received support from a researcher suggesting a response when a child had tried several times without success and looked puzzled or begged for help. In 69 cases, they answered don't know, meaning that this question will reoccur at a later stage. This result showed us that it was at least technically possible to create an interaction between the game, the robot, and the children.

Interestingly, during this phase, the children became more focused on their interactions with the game, trying to play it well. In some cases, this led them to completely ignore the robot's questions. Although we can see this as a sign of the potential effectiveness of our first phase, where we demystified the robot, we of course intend the children to become engaged in the interaction with both the game and the robot tutee in order to learn arithmetic. It is thus possible that some novelty effect was created through the introduction of the game. It became clear that the game on the interactive board, and the robot standing beside the child, were two simultaneous sources calling (and competing) for the child's attention. Some children focused only on the robot tutee, some only on the game (and had to be reminded to respond to the robot tutee), and the rest switched between the two sources fairly spontaneously.

After the interaction with Pepper and the game, one of the researchers sat down with each group to engage them in a semi-structured reflection discussion. During this discussion, many children highlighted how Pepper had developed compared to the robot they had met the first time. They often acknowledged that Pepper's more advanced cognitive abilities were related to programming: "The programming is really good so yes I think that he or she, I do not know, but now it is smart programmed, good and all, well it is smart." (Fourth grader). However, they were also aware that the robot's smart behavior was not something magical: "But basically so is it not smarter it is not smarter than us. It is not smarter because we create the robot so we are smarter" (Fourth grader). While the children had previously mainly defined robots in terms of possessing human attributes, such as eyes, legs, heads and fingers, the distinction between machines and robots had now blurred. Some children now considered vacuum cleaners and washing machines to be robots too, while other still maintained that robots should have some human qualities. The reflection discussion confirmed our observation that the children's attention had shifted from the robot to the newly introduced arithmetic game. Although the general consensus still was that meeting Pepper was exciting and fun, when asked about the most positive aspect of the workshops, the children often mentioned the game. They also commented positively about learning mathematics using the game, or using the game together with Pepper: "But then again we also learn from teaching so that is also really good, everyone learns" (Fourth grader).

Similar to the first phase, we sent out a post-workshop questionnaire to the children focusing on their experiences with the game-playing robot tutee. We wanted to understand whether they indeed had experienced their interaction with the robot as it having the role as a tutee. Furthermore, we wanted to know what they thought of Pepper's questions and Pepper's behavior and in what direction they would Pepper to develop. We received answers from 67 children, which were analyzed both quantitatively and qualitatively. The detailed results of this analysis are presented in another paper [36]. However, the results indicated that the children indeed experienced the interaction with Pepper as if they were teaching it and not vice versa. While most of them indicated that they were able to detect and identify what questions Pepper had asked (which is not self-evident considering the speech synthesis technology used in robots), the answers were slightly less positive as to whether they had understood the questions. In the future, they wanted Pepper to ask more mathematics-related questions, e.g., more challenging questions such as "what is 9x6?" or "what is 100x100?"; more social questions, by asking for the child's name and age, how they liked the game or mathematics in general, and giving them the opportunity to choose a game; and pedagogical questions, such as questioning their game moves (e.g., "Are you sure?") and offering more help and hints. Finally, we studied what the children thought about Pepper's current game-playing behavior and sought suggestions for improvement. While about half of the children did not have any suggestions for improvement, the other half specifically mentioned problems with the eligibility of Pepper's speech, and technical problems like becoming unfocused (this can happen when the sensors get too much input), shutting down, or that it stopped talking. Only one child commented that Pepper asked too many questions, and one child expressed some disappointment that Pepper had limited functionality. They also wanted Pepper to be able to choose a card by itself by interacting with the interactive board. They also pointed out that Pepper should switch gaze between the player and the game and move more (also different kinds of movements). Finally, they thought that Pepper should talk more about other subjects than mathematics, but also ask even more mathematics-related questions.

5 Discussion and Further Steps

The two phases of our co-design approach to create a robot tutee in a school context described above, are part of a longer engagement with two schools. Such long engagement with the same schools and children is important to get beyond the novelty effect of interacting with a robot. In these first two phases, we focused mainly on demystifying the robot and introducing the children to the idea of a robot tutee. The results of the first phase clearly indicated that our intention to demystify robots by showing children how they must be programmed before they can do anything, was successful. While this demystification resulted in a slight disappointment for some children, others were enthusiastic about the opportunity to help define the robot's behavior. The experiences during the first phase helped to create a calmer atmosphere in the second phase, where we aimed to introduce the children to the idea of Pepper functioning as a robot tutee. Indeed, the attention for some of the children even shifted away from the robot towards the game as a novelty. The second phase also gave us an indication that the setup with the robot tutee might work, even though there are still several technical aspects that need to be improved.

After having worked with the children during these two phases, we are confident that we now will be able to involve them in the next phases of the project where we aim to refine the behavior of the robot tutee and eventually give them the possibility to use the robot and the game in their classroom. While not addressed in his paper, we also perform co-design activities with the teachers in parallel. We have, e.g., performed workshops with the teachers where they were asked to write dialogues for the robot tutee focusing more directly on the mathematical content.

As a take-away for other projects aiming to co-design educational robots with children, we would like to suggest the importance of a demystifying phase. This can probably be done in many different ways, and probably with robots in other roles (e.g., as tutors or companions) and domains (e.g., in healthcare or entertainment), but we think our choice to show how robots are programmed by starting with a relatively blank slate was a successful approach. The purpose of the demystifying phase is to allow stakeholders unfamiliar to a certain technology to become aware of its potentials and limitations, and thus be better prepared to contribute in the successive co-design phases. This take-away message may apply for co-design processes concerning other educational technologies, particularly those that may be unfamiliar to children (e.g., virtual agents).

Acknowledgments. We would like to thank the children and teachers from the two schools working as co-design partners in our project. This work was partially supported by the Marcus and Amalia Wallenberg Foundation and was funded in part by the project START, which is a collaboration between University West and the University of Gothenburg. The authors are solely responsible for the content of this publication. It does not represent the opinion of the Marcus and Amalia Wallenberg Foundation, and the Marcus and Amalia Wallenberg Foundation is not responsible for any use that might be made of data appearing therein.

References

1. Druin, A., *The role of children in the design of new technology*. Behaviour & Information Technology, 2002. **21**: p. 1-25.
2. Druin, A., ed. *The design of children's technology*. 1999, Morgan Kaufmann Publishers: San Francisco, CA.
3. Druin, A. *Cooperative inquiry: developing new technologies for children with children*. in *SIGCHI conference on Human Factors in Computing Systems, CHI '99*. 1999. ACM.
4. Axelsen, L.V., L. Mygind, and P. Bentsen, *Designing with Children - A Participatory Design Framework for Developing Interactive Exhibitions*. International Journal of the Inclusive Museum, 2014. **7**.
5. Danielsson, K. and C. Wiberg, *Participatory design of learning media: Designing educational computer games with and for teenagers*. Interactive Technology and Smart Education, 2006. **3**(4): p. 275-291.
6. Lee, H.R., et al., *Steps Toward Participatory Design of Social Robots: Mutual Learning with Older Adults with Depression*, in *Proceedings of HRI'17*. 2017, ACM: Vienna, Austria. p. 244-253.
7. Šabanović, S., et al., *Participatory design of socially assistive robots for preventive patient-centered healthcare* in *IROS Workshop on Assistive Robotics for Individuals with Disabilities*. 2014, IEEE: Chicago, IL.
8. Šabanović, S., S. Reeder, and B. Kechavarzi, *Designing Robots in the Wild: In situ Prototype Evaluation for a Break Management Robot*. Journal of Human-Robot Interaction, 2014. **3**(1).
9. Reich-Stiebert, N., F. Eyssel, and C. Hohnemann, *Involve the user! Changing attitudes toward robots by user participation in a robot prototyping process* Computers in Human Behavior, 2018.
10. Pareto, L., *Robot as tutee*, in *Robotics in Education*. 2017, Springer. p. 271-277.

11. Roscoe, R.D. and M.T. Chi, *Understanding Tutor Learning: Knowledge-Building and Knowledge-Telling in Peer Tutors' Explanations and Questions*. Review of Educational Research, 2007. **77**(4): p. 534–574.
12. Börjesson, P., et al., *Teachers' Expected and Perceived Gains of Participation in Classroom Based Design Activities*, in *Proceedings of CHI'19*. 2019, ACM: Glasgow, Scotland Uk. p. 1-9.
13. Druin, A., et al. *Designing PETS: A Personal Electronic Teller of Stories*. in *CHI '99*. 1999. ACM Press.
14. Obaid, M., et al. *Designing Robotic Teaching Assistants: Interaction Design Students' and Children's Views'*. in *ICSR '15*. 2015. Paris, France.
15. Obaid, M., et al., *Developing a prototyping method for involving children in the design of classroom robots*. International Journal of Social Robotics, 2018. **10**(2): p. 279-291.
16. Woods, S., *Exploring the design space of robots: Children's perspectives*. Interacting with Computers, 2006. **18**: p. 1390-1418.
17. Arnold, L., J. Yip, and J.L. Kung. *Co-designing with children: An approach to social robot design*. in *HRI'16*. 2016. Christchurch, New Zealand.
18. Bertel, L.B., D.M. Rasmussen, and E. Christiansen. *Robots for Real: Developing a Participatory Design Framework for Implementing Educational Robots in Real-World Learning Environments*. 2013. Berlin, Heidelberg: Springer Berlin Heidelberg.
19. Rohrbeck, C., et al., *Peer-assisted learning interventions with elementary school students: A meta-analytic review*. Vol. 95. 2003. 240-257.
20. Blair, K., et al., *Pedagogical Agents for Learning by Teaching: Teachable Agents*. Educational Technology, 2007.
21. Chin, D.B., et al., *Preparing students for future learning with Teachable Agents*. Education Tech Research Dev, 2010. **58**(6): p. 649-669.
22. Walker, E., et al. *The Effects of Physical form and Embodied Action in a Teachable Robot for Geometry Learning*. in *(ICALT'16*. 2016.
23. Dautenhahn, K., *Socially intelligent robots: dimensions of human–robot interaction*. Philosophical Transactions of the Royal Society B: Biological Sciences, 2007. **362**(1480): p. 679-704.
24. Reeves, B. and C. Nass, *The Media Equation : How People Treat Computers, Television, and New Media like Real People and Places*. CSLI Lecture Notes. 1996: The Center for the Study of Language and Information Publications.
25. Thomaz, A.L. and C. Breazeal, *Teachable robots: Understanding human teaching behavior to build more effective robot learners*. Artificial Intelligence, 2008. **172**(6-7): p. 716-737.
26. Zhang, J. and A.J.C. Sharkey, *It's not all written on the robot's face*. Robotics and Autonomous Systems, 2012. **60**(11): p. 1449-1456.
27. Pareto, L., et al., *A teachable-agent-based game affording collaboration and competition: Evaluating math comprehension and motivation*. Educational Technology Research and Development, 2012. **60**(5): p. 723-751.
28. Pareto, L., *A teachable agent game engaging primary school children to learn arithmetic concepts and reasoning*. International Journal of Artificial Intelligence in Education, 2014. **24**(3): p. 251-283.
29. Bratteteig, T. and G. Verne, *Does AI make PD obsolete?: exploring challenges from artificial intelligence to participatory design*. 2018. p. 8:1-8:5.
30. Kam, M., et al., *Practical considerations for participatory design with rural school children in underdeveloped regions: early reflections from the field*, in *Proceedings of the 2006 conference on Interaction design and children*. 2006, ACM: Tampere, Finland. p. 25-32.

31. Bossen, C., C. Dindler, and O.S. Iversen, *User gains and PD aims: assessment from a participatory design project.*, in *11th Biennial Participatory Design Conference (PDC '10).* 2010. p. 141-150.
32. Iivari, N. and M. Kinnula, *Inclusive or Inflexible: a Critical Analysis of the School Context in Supporting Children's Genuine Participation*, in *NordiCHI'16*. 2016, ACM: Gothenburg, Sweden. p. 1-10.
33. Barendregt, W., et al., *Legitimate Participation in the Classroom Context: Adding Learning Goals to Participatory Design*, in *IDC'16*. 2016, ACM: Manchester, United Kingdom. p. 167-174.
34. Serholt, S., *Interactions with an Empathic Robot Tutor in Education: Students' Perceptions Three Years Later, in* in *Artificial Intelligence and Inclusive Education: Speculative Futures and Emerging Practices*, J. Knox, Y. Wang, and M. Gallagher, Editors. 2019, Springer Singapore: Singapore.
35. Braun, V. and V. Clarke, *Using thematic analysis in psychology*. Qualitative Research in Psychology, 2006. **3**(2): p. 77-101.
36. Pareto, L., Ekström, S., Barendregt, W., Serholt, S., and Kiesewetter, S. *Augmenting Game-Based Learning with a Robot Tutee*. in *13th European Conference on Games Based Learning (ECGBL)*. 2019. Odense, Denmark: Academic Conferences and Publishing International Limited.