EU FP7 SQUIRREL
610532
Feb 1, 2014 (48months)

D2.1:
Report on user behavior

Vicky Charisi, Manja Lohse, Cristina Zaga, and Vanessa Evers

University of Twente, Enschede
⟨v.charisi,m.lohse,c.zaga,v.evers@utwente.nl⟩

Due date of deliverable:	Month 12 (January 2015)
Actual submission date:	13 March 2015
Lead partner:	University of Twente
Revision:	final
Dissemination level:	PU

This deliverable presents the first results of WP 2 (Robust User Interaction in Crowded Environments). It addresses mainly T2.1 (User Behavior Analysis and Scenario Definition) based on a contextual analysis on children's collaborative play (without robot) and a first user study on children's engagement with a robot. It also includes a first corpus of annotated data from the first user study.

The main goal of the contextual analysis was to identify processes and patterns of how children deal with clutter in their social interactions. We researched what non-verbal behaviors appeared in children's small-group collaborative activities and how each child's behaviors affected the other children in the group.

As engagement is one of the main aspects that contribute to initiate, sustain and maintain child-robot interactions, it was also the focus of the first study with children and a robot. We investigated the effect of a robot's social characters expressing two different styles of interaction (i.e., peer and tutor) on children-task engagement and task performance. The annotated corpus of this study (i.e., ELAN files and videos) is part of this deliverable and can be accessed by the project partners.

The results of both studies inform us about how the children get engaged in the task and how a robot could help this process. This work sets the frame for our on-going research on robot behavior design in T2.4. At the same time, both studies inform the SQUIRREL scenarios and the work of other partners in the project such as the ones working on perception and planning.

1 Summary **4**

2 Contextual Analysis **5**
2.1 Theoretical background 5
2.1.1 Constructivism and child's development 5
2.1.2 Socio-cultural theory 6
2.1.3 Creative thinking 7
2.1.4 Summary 7
2.2 Related work 8
2.2.1 Collaborative activities 8
2.2.2 Play and engagement 8
2.3 Research questions 9
2.4 Method 9
2.4.1 Participants 10
2.4.2 Ethical considerations 11
2.4.3 Task design 11
2.4.4 Setup 13
2.4.5 Procedure 14
2.4.6 Pilot study 17
2.5 Measures and data analysis 18
2.5.1 Annotation Scheme 18
2.5.2 Task Performance 19
2.5.3 Actions 20
2.5.4 Positioning 22
2.6 Results 25
2.6.1 Task Performance 25
2.6.2 Actions 25
2.6.3 Positioning 29
2.7 Discussion and Design Implications 45

3 First user study on child-robot interaction **49**
3.1 Introduction 49
3.2 Task Engagement 50
3.3 Related Work 51
3.4 Hypotheses 53
3.5 Method 54
3.5.1 Robot and WoZ 55
3.5.2 Behavior design 56
3.5.3 Task scenario 58
3.6 Setup and Procedure 60
3.7 Data analysis 62
3.8 Results 65
3.8.1 Cognitive attributes of task engagement: focus of attention 65

3.8.2 Affective attributes of engagement: enjoyment 66
3.8.3 Behavioral attributes of task engagement: task completion 66
3.8.4 Behavioral attribute of task engagement: efficiency . . 67
3.8.5 Children's perception of the robot: peer or tutor . . . 67
3.9 Discussion 67
3.10 Limitations and future work 69

4 Corpus of data from first user study **71**
4.1 Use of the corpus: ethical consideration 71
4.2 Videos 72
4.3 ELAN Files 73

1 Summary

This deliverable presents the first results of WP 2 (Robust User Interaction in Crowded Environments). It addresses mainly T2.1 (User Behavior Analysis and Scenario Definition) based on a contextual analysis on children's collaborative play (without robot) (see Section 2) and a first user study on children's engagement with a robot (see Section 3). It also includes a first corpus of annotated data from the first user study (see Section 4).

The main goal of the contextual analysis was to identify how children deal with clutter in their social interactions. We researched what non-verbal behaviors appeared in children's small-group collaborative activities and how each child's behaviors affected the other children in the group. To address these points, we observed 24 children, aged 5-8, in groups of 3. The groups assembled structures from magnetic wooden blocks. We analyzed the video recordings for children's actions, overlap between these actions, and spatial behavior in the task. The results inform us about how the children get engaged in the task and how a robot could help this process.

As engagement is one of the main aspects that contribute to initiate, sustain and maintain child-robot interactions, it was also the focus of the first study with children and a robot. Whilst some efforts to identify the features of human-robot task engagement have been reported in the literature, little is known about the patterns of children-robot task engagement. In particular, it is not yet well understood which style of interaction a robot should have to successfully maintain and sustain the task engagement of more than one child which is very relevant for SQUIRREL. We investigated the effect of a robot's social characters expressing two different styles of interaction (i.e., peer and tutor) on children-task engagement and task performance. 10 pairs of children, aged 6 to 9, participated in this study. Our results showed higher task effectiveness and performance in the peer condition (where the robot behaved as a friend) at least in more difficult tasks. The annotated corpus of this study (i.e., ELAN files and videos) is part of this deliverable and can be accessed by the project partners.

This work sets the frame for our on-going research on robot behavior design in T2.4. At the same time, both studies inform the SQUIRREL scenarios and the work of other partners in the project such as the ones working on perception and planning.

2 Contextual Analysis

In daily life, children have to deal with different situations, which function as a context for them to develop and learn. Often, they play in crowded environments with other children or adults as well as with the clutter, which implicitly results in their development. While they engage in tasks or games, children exhibit a variety of social behaviors. In order for us to design an appropriate context for the children within the SQUIRREL project, we need to develop a deep understanding of how children deal with clutter in their daily life. Thus, the main goal of this contextual analysis was to identify how children deal with clutter in their social interactions. We call this study a contextual analysis as a reference to a method of sociological analysis [23]. Based on this method we observe children in their natural context (here: a school environment) and interpret their behaviors in the context of the interaction in the group.

In line with the SQUIRREL project goal to clean clutter bit by bit, we took the decision to focus the contextual analysis on children's spatial awareness and sorting skills. The task also included opportunities for creative construction. In this way, we tried to transform the activity of tidying up, which is often considered as a routine activity in children's life, into an opportunity for their creative expression through play.

We based our work on fundamental psychological theories for children's development. We explain basic aspects of those theories in Section 2.1. Relevant current studies that are based on and have expanded those theories are reported in Section 2.2. Following this, we describe the research questions (see Section 2.3), the methods (see Section 2.4), the setup (see Section 2.5) and the findings (see Section 2.6) of the contextual analysis. At the end of the section we present a discussion and design implications (see Section 2.7).

2.1 Theoretical background

The theoretical framework that has been used for understanding and analyzing children's behaviors is underpinned (i) by the notion of constructivism, (ii) by aspects of the socio-cultural theory and (iii) by the theories on children's creative thinking processes and play.

2.1.1 Constructivism and child's development

Piaget [52] outlined children's cognitive development as a progressive reorganization of mental processes, which underpin the construction of the increasing number and complexity of children's schemata. A schema is defined as a set of linked mental representations of the surrounding world and it is formed and grown through children's development. This occurs, according to Piaget, through a sequence of processes. It starts with the use of an existing schema, which needs to be changed in order to deal with a new situation and, eventually, with all similar situations.

Piaget conceived children's cognitive development in the form of four stages, which relate to specific age-groups, namely (i) sensorimotor stage, (ii) preoperational stage, (iii) concrete operational stage and (iv) formal operational stage. In line with the SQUIRREL project, the participant in this research are children that attend primary education and belong to the concrete operational stage, which refer to children from 6-7 years old up to 10-11 years old. In this stage children develop their logical thinking and are able to handle more abstract concepts. They become less ego-centric and are able to realize and share their thoughts and feelings. Thus, in this study, we expected from children to be able to understand the rules of an activity and to develop their spatial reasoning and sorting skills. Also, according to this taxonomy, children in this age are able to play collaboratively and interact with each other during an activity.

2.1.2 Socio-cultural theory

Socio-cultural theory prioritizes the importance of social interactions in child's learning processes. Vygotsky [75] discussed that children learn in a better and more efficient way if they interact with other children or adults. In this way, they have more opportunities to verbalize their thoughts, which makes them understand the world around them in a better way.

In addition to social interactions, Vygotsky highlighted the important role of 'tools' in children's learning processes and development. Tools mediate any human activity and interaction with the physical world and consequently their features and design affect our thinking [29]. In the context of children's activities appropriately designed tools may affect children's development in a better way.

Through their social interaction and their interaction with specific tools, children have the chance to exploit and expand their Zone of Proximal Development.The concept of Zone of Proximal Development (ZPD) has been introduced by Vygotsky to describe the level of their potential development, in which learning takes place. This concept refers to the time when child's spontaneous concepts meet the systematic reasoning through her interaction with elements of the environment and her peers or other adults [20].

In accordance with this theory we took into consideration both opportunities for social interactions as well as the possibilities that specific tools encourage engagement and development. Hence, we aimed to design an interactive task that is supported by engaging tools.

2.1.3 Creative thinking

For this contextual analysis, elements of theories of children's creative thinking processes have been taken into consideration for the design of a developmentally appropriate task that is rich in opportunities for learning and creation. An innate characteristic that facilitates children's development is their need to explore their surrounding world and to adapt to it by using their sensory skills. Children's inherent motivation to explore their surrounding world has been related to their imaginative and creative approaches to explore, understand and act upon their environment through playful activities [39], [56]. Since creativity is a human characteristic that underpins every innovation, scientific discovery or piece of music and art, developmental psychologists emphasize the importance of its understanding and support in children's daily life [80].

Torrance [71] emphasized the importance for children to develop creatively, giving children a chance to learn, think and act in a variety of ways that may also contain mistakes. During such a process, children learn by means such as questioning, experimenting, tinkering and often by aimless play. This approach has been recently seen in relation to children's playfulness [39] and possibility thinking [21], which involves curiosity-driven exploration in childhood.

Wallas [77] offered an analytic framework of a relatively valid way to organize the complexities, which are involved in the process of creative thinking, around four fundamental stages, namely (i) preparation, (ii) incubation, (iii) illumination and (iv) verification. This approach of children's processes emphasizes to intentional/deliberate actions towards the completion of a task.

More recently, creative thinking has been considered with regard to its social nature. Holzman [32], for example, highlights the correlation of ZPD expansion with creative thinking. Researchers suggest that creativity happens in the interaction between a persons thoughts and a socio-cultural context [4]. This approach of creative thinking is interpreted as being fundamentally underpinned by the concepts of social constructivism and Vygotskian socio-cultural theory, which have been discussed in the previous paragraphs.

For the design of this study we took into consideration the fundamental aspects of these theories of children's creative thinking processes. The design of the task, that children went through, was mainly based on the basic stages of the creative process, namely exploration, elaboration, and creation.

2.1.4 Summary

The combination of the constructivist approach with Vygotsky's socio-cultural theory in the context of open-ended activities that allow for collaboration and creative thinking has been considered as the theoretical framework for this study. Papert [50] emphasized the importance of children's flexibility of knowledge under construction. In this approach children are allowed to take a risk, to make mistakes and to mess around with risky ideas. Learning occurs when they try to make sense of and deal with the messy, crowded status of their ideas.

2.2 Related work

In this section we focus on recent studies and related work that have informed two core aspects of this study, namely (i) children's collaborative activities and (ii) the contextualization of their creative thinking with playful activities and their engagement.

2.2.1 Collaborative activities

Empirical and neuroscientific research has shown that children develop in a more effective and holistic way when they interact with other people (e.g., [10]). This interaction may have different forms, one of which is collaboration. Collaboration occurs when two or more people work together towards a common goal [59]. Rogoff defines collaboration as the mutual involvement and participation in shared endeavors, which may or may not strive to promote cognitive development. In collaborative activities, children share and discuss the actions they take and their products are conceived as a result of these shared actions [63].

More recently, Tomasello [70] highlighted a distinct element of collaborative activities: collective intentionality. When all members of the team are aware of their common goal, they tend to work more effectively together. However, collective intentionality appears to be dynamic and evolving within the members of the same group. Accordingly, Wing-yi Cheng et al. [81] argue that what is more important for children's collective intentionality and efficacy is related to group processes rather than the homogeneity or inhomogeneity of the group. In their big scale research (N=1921 students), they found that both high achievers as well as low achievers may benefit from group work.

As the SQUIRREL project focuses on groups of children interacting with a robot, in this research, we emphasize the importance of groups' collective intentionality, since this contributes more effort to the group endeavor, which results in greater group and individual accomplishments [5].

2.2.2 Play and engagement

Another crucial characteristic of children's development is their natural need for play, which mediates and partially determines the degree of their task engagement. Play has been considered as the most natural way for a child to develop [28]. According to Froebel, play provides all the means for a child's holistic development, which includes her cognitive, social, emotional and physical development. Contextualizing this study with playful activities that might be more meaningful, creative and enjoyable for children makes it more likely to achieve a higher degree of children's task engagement [51], [58].

In the literature, engagement has been described as a quality of human experience, which is characterized by attributes of challenge, positive affect, endurability, attention, to mention a few [48]. Wang and Degol [78] highlight the importance of considering the multidimensionality of engagement and its dynamic character.

In this study, we focus on children's behaviors that indicate their task engagement and their social engagement in a playful activity.

2.3 Research questions

The primary goal of this contextual analysis is to understand the behavior of children in their everyday school environment and the ways that they collaborate in order to accomplish their task goals. The results of this contextual analysis will inform the next research steps for the SQUIRREL project regarding robot's behaviors to optimize children's interaction with it. Thus, an understanding is needed of children's behaviors in a collaborative activity. As the SQUIRREL robot will mainly use non-verbal behaviors itself, we here focus on these rather than on speech.

For this analysis we take into account both individual behaviors as well as interrelations in the group. In accordance with the previously mentioned goals as well as with findings from relevant research studies, we formed the research question of this contextual analysis:

- What kind of relevant non-verbal behaviors appear in children's small-group collaborative activities?
- How do the non-verbal behaviors affect task engagement of all children in a group?
- How do individuals' behaviors combine to form different types of childrens collaboration?

2.4 Method

The most suitable method for us to collect data for this contextual analysis was participatory observation in a naturalistic environment. Observation is a widely used method for research studies that focus on children's behaviors, since it can provide a systematic description of events and behaviors in the social setting chosen for study (e.g., [17], [44]). The observation in our case

Group	*Child 1*	*Child 2*	*Child 3*
1	6 / M	7 / F	8 / M
2	7 / F	7 /M	7 / F
3	7 / M	6 / M	7 / F
4	8 / M	8 / F	8 / M
5	7 / M	6 / F	7 / M
6	6 / M	6 / F	6 / M
7	6 / M	6 / M	6 / M
8	9 / M	6/ F	5 / F

Table 2.1: Demographics (age / gender) of each child in each group

was participatory, i.e., one researcher functioned as facilitator, e.g., explaining tasks to children. We considered this necessary given the age group of our participants in order to decrease uncertainties and stress. However, the facilitator intervened as little as possible, in order to let children regulate their own ways and take their own decisions of solving the task. This is in accordance with the fundamental theories that have been described in Sections 2.1 and 2.2.

Data collection took place in a public primary school in the Netherlands. A specific system (Montessori) underpins the pedagogical approaches of this school, which support learning by doing with the use of sensorimotor collaborative activities. As a result, children in this school were familiar with working in small groups using different kinds of physical tools. For the purposes of this research, we recruited participants from age 5-8 (see Section 2.4.1). All the sessions were video-taped and audio-recorded (see Section 2.4.4). The recordings were then annotated in the software ELAN. In total, we collected data from 8 sessions that took approximately 30 minutes each. In the whole process, including the task design (see Section 2.4.3) and the procedure (see Section 2.4.5), we paid specific attention to ethical considerations as laid out in Section 2.4.2. We also conducted a pilot study (see Section 2.4.6) to practice the process, to test our measures (see Section 2.5) and to make sure that the participation was as little stressful and as much fun for the children as possible.

2.4.1 Participants

The participants were students of early primary education. In total 24 children (10 girls and 13 boys) aged 5-8 (M=6.95 and SD=0.95) were assigned to groups of 3 (see Table 2.1). We decided to investigate children's behaviors in such groups that are considered to be the smallest possible social group (Stangor, 2004). We believed that this group size would trigger collaborative activities and different collaborative constellations.

The children were assigned according to their age, trying to have children with approximately the same age in the same group. All the children participated in the same tasks. Prior to participation the parents had to fill in a consent form as explained in depth below.

2.4.2 Ethical considerations

During the research design, we ensured that this study conformed to the ethical procedures, which are in compliance with the fundamental ethical principles that are suggested by the Code of Human Research Ethics [13] and the Data Protection Act as it appears at the British Education Research Association [8]. Before the data collection, we received the initial permission by the head of the school and we followed up with a discussion with the teachers of the school about the aim of the research. Subsequently, we sent to the parents and carers of the children a form, in order to obtain their approval for their child to participate in the study. Finally, the children who had parental permission to participate were asked whether they wanted to take part in the study. In accordance with research that advocate the importance of listening to children's voices, e.g. [16], the researcher informed children that their participation in the research was voluntary and that they had the right to withdraw from the study at any time for any or no reason. The best interests of the children were the primary consideration throughout the research study. There was no predictable detriment arising from the process of the research. Confidential and anonymous treatment of participant's data was considered the norm in this research. Although the issues of confidentiality and anonymity are particularly critical for all research, children's vulnerability heightened the importance of protecting their identities. It was made clear that any resultant data from the study would be anonymized and the participants would not be identified at any point. Participants' privacy was protected by avoiding any undue intrusion into their personal affairs [2]. They were also informed that data were to be kept safely on a secure server and password protected.

2.4.3 Task design

Since the goal of this research does not predetermine a specific task, we decided that a flexible and open-ended environment, which supports children's choices for action, would be the ideal playful activity. The task we decided for was 3D constructions with magnetic wooden blocks. This task is in line with the idea of clearing clutter because its basic component is sorting blocks. We chose the specific magnetic wooden blocks, because they have features that are developmentally appropriate for children of this age-group. The blocks have a variety of shapes, colors, and directions of the magnets that allow to construct many shapes with different levels of difficulty. For our task, all the shapes depicted animals because children are familiar with them and this theme allows for a lot of variation.

Playing with wooden blocks may have multiple benefits for children. Activities with building blocks have been mainly related to the development of spatial awareness and have been reported as crucial for the development of reasoning skills, the success in mathematics and science, the abilities of navigation, estimation and balance [72], [76]. These skills are highly relevant for the SQUIRREL scenario of clearing clutter bit by bit.

For the purposes of this research, we distinguish two kinds of activities with building blocks: (i) structured play and (ii) unstructured play. In the first case, children are given a model (image) that they have to build. In the second case, children engage in open-ended block play. In both cases, as Verdine et al. [74] reported, inherent geometric properties of blocks encourage the development of children's ability to produce complex relations and analytical skills of spatial representations. These skills are relevant for children as well as for the robot collaborating in the SQUIRREL scenario.

The structure of the activity was based on the initial theoretical framework about the process of creativity and on existing models of young children's creative thinking processes (see Section 2.1). The main features that frame the design of this activity are: (i) Exploration, (ii) Guided Construction, and (iii) Creative Construction. Thus, we decided for three tasks, each of them focusing on one of these aspects.

For the first two tasks, the shapes that the children had to construct were presented to them in pictures. Each shape followed some rules of coherency that the children had to discover. For example, each part of the animal body (head, body, legs and tail) consisted of equal shapes (cubes, quadrilaterals, triangles, or rectangles) and/or colors (green, blue, brown or pink). In the following the three tasks are described in some more depth.

Exploration (Task 1) In the exploration phase the children were free to play with the blocks and to try out how to assemble a simple shape, provided to them in an image, without time limit. The purpose of this task was twofold: on the one hand, we gave the children the opportunity to explore the blocks and the possibilities of their combinations; on the other hand children explored and formulated thinking strategies in order to replicate the specific structure that were depicted in the image.

Guided Construction (Task 2) In this task again the children were asked to replicate the structures that were depicted in images. In total, six different images were designed for children with an increasing degree of difficulty. The difficulty depended on the number of the blocks that were needed to replicate the structure and on the complexity of block combinations. Children were allowed to choose the degree of difficulty of the structures. However, they

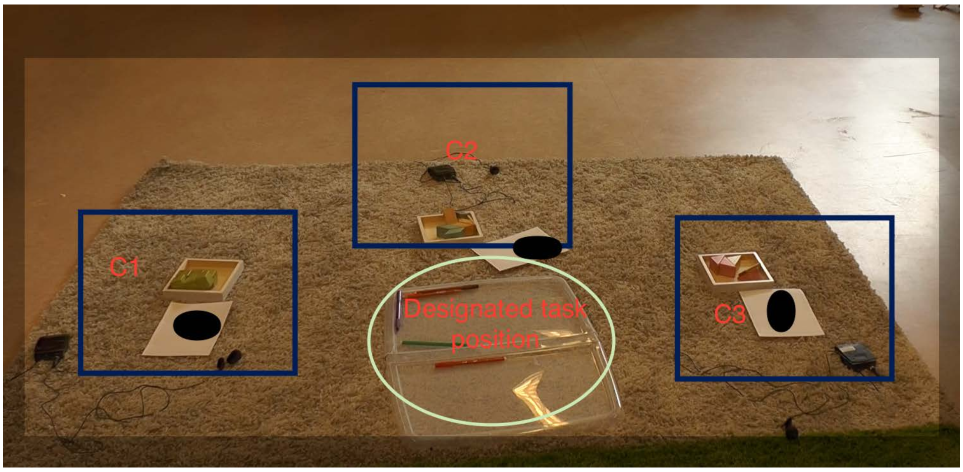


Figure 2.1: Setup of the experimental area

.

were not allowed to see each picture in advance.

Creative Construction (Task 3) The third task was designed to let children's creativity unfold while they used their sorting skills and spatial reasoning that was developed in the previous tasks. Children had to build an animal but were completely free in what and how to build. The facilitator emphasized that we expected the children to build one construction together.

2.4.4 Setup

We conducted the research in a local primary school in the Netherlands. The head of the school gave us access to a room that is normally used for after-school activities. It was familiar to the children and the furniture and the decoration were suitable to make children feel at ease.

We decided to conduct the research on the floor of this room, avoiding tables and chairs (see Figure 2.1). In this way, we tried to give a more playful air to the activity and allow children to move within the playground. However, we put two carpets in order to implicitly define the barriers of the playground. At the two sides of the playground we put books in order to define the playground in a better way and to protect children from any distraction by the researchers.

We used one small table and a chair, outside the playground in distance of approximately 2m from the carpet, for storing the relevant material for the study (e.g. pictures, worksheets) and for the facilitator to sit, while the children were engaged with the task. Also, we used two tables, out of vision of the participating children, to put all the technical material (e.g.,

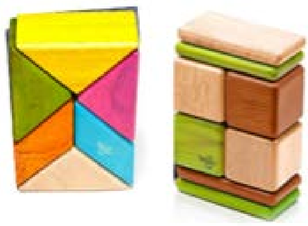


Figure 2.2: Magnetic wooden blocks used in the study.

.

laptops). We placed a plastic frame in the center of the playground in order to support childrens collaboration, since in this way we implicitly guided children to work in the same place. Also, we placed three different boxes next to the position of each child, in which we put the magnetic blocks.

In each box, we placed a limited number of blocks that the children needed to build a specific shape. In total, we used twelve cubes, twelve right-angle parallelograms, twelve triangles and six parallelograms (see Figure 2.2). They were split between children in a way that blocks from each child were needed to complete a shape.

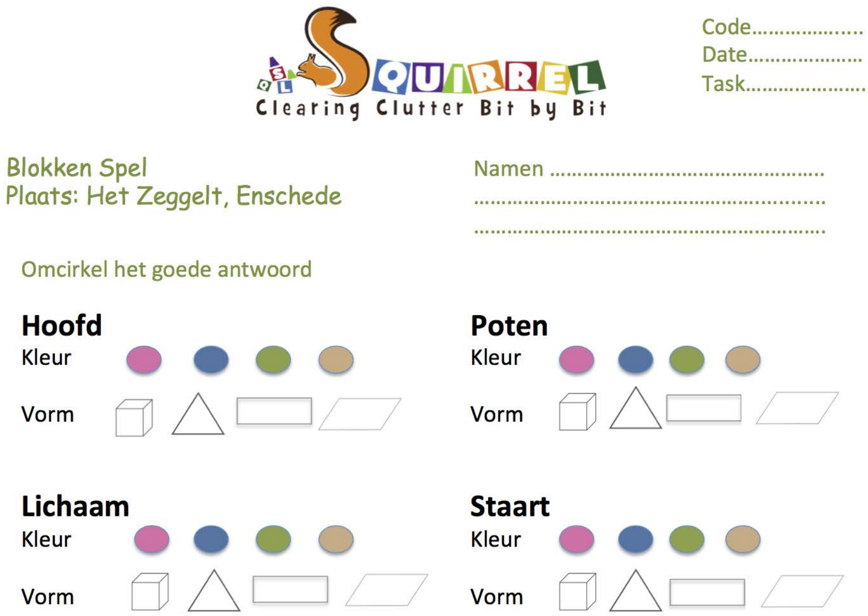
To give children the opportunity to reflect on their work, we designed a worksheet that they had to complete after each construction (see Figure 2.3).

For data collection purposes, we video-recorded all sessions. In total, three cameras recorded each child's behaviors. The central camera recorded the overview of the activity. A camera to the right recorded the child at the left and vice versa.

Three members of the research team were present during all the sessions. However, they only took care of the technical setup. We collaborated with a research assistant, who facilitated children to go through the activities. The facilitator was a Dutch-speaking female with background in educational sciences. She was given specific instructions to interact with the children during the activity, with the main rule being to intervene as little as possible.

2.4.5 Procedure

When the children entered the room, they were guided to their positions on the play mat (see Section 2.4.4). They were equipped with close talk microphones and asked to assent to participation. If they assented, we turned on the cameras to record the interaction. Thereafter, the facilitator asked them about their names and explained the activity. After this brief introduction, the children started to work on the first task.



SQUIRREL
Clearing Clutter Bit by Bit

Code..................
Date..................
Task..................

Blokken Spel
Plaats: Het Zeggelt, Enschede

Namen ....
....
....

Omcirkel het goede antwoord

Hoofd
Kleur
Vorm

Poten
Kleur
Vorm

Lichaam
Kleur
Vorm

Staart
Kleur
Vorm

Figure 2.3: Worksheet used in the study asking the children to define the colors and shapes of the parts of the animals' bodies.

Each of the three tasks introduced above was designed to have a fixed duration and complexity level. However, during the actual data collection, we were not absolutely strict as for the duration of the tasks, allowing children to take some more minutes if needed, particularly in the exploration task. This first task contained the following steps that took about five minutes to complete:

- A picture of a shape was given to the children (see Figure 2.4)

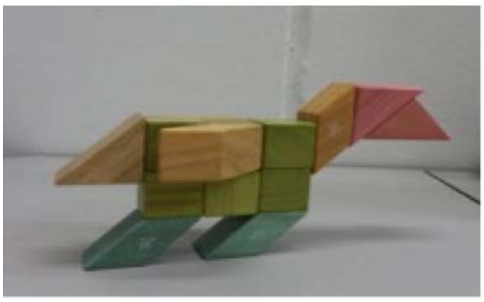


Figure 2.4: Image provided to the children in the first task (exploration).

- The facilitator explained to the children that they had to build the

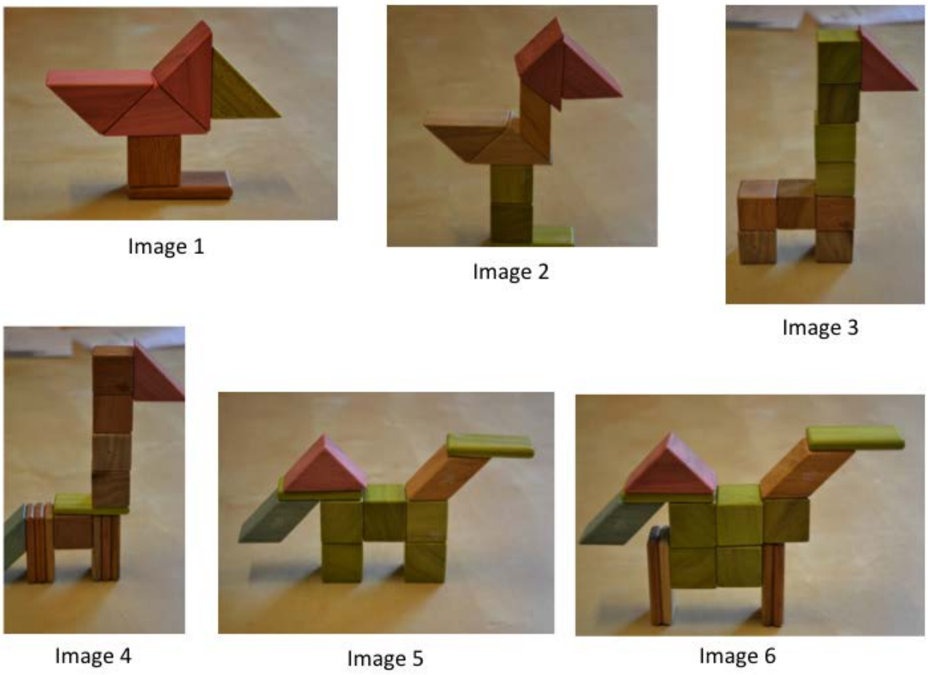


Figure 2.5: Images provided to the children in the second task (guided construction).

structure of the image with the wooden blocks. Only 18 cubes were given to the children, which include the cubes that were needed for the construction of the specific image plus 30% of cubes that were not actually needed for the construction. The cubes were distributed among the children in accordance to their shape, in order to maximize the possibility of collaboration.

- Children completed the task and the worksheet.
- The facilitator made sure that the children had built the structure and completed the worksheet. She gave positive feedback about the construction. Then, she allowed the children to deconstruct the model.

The second task (guided construction) consisted of six structures that the children could built (see Figure 2.5). The children were asked to complete as many structures as possible within 10 minutes. The structures were of gradual difficulty. The difficulty was related to the number of the blocks as well as to the complexity of their combinations. However, one rule was common to all the constructions in order to create coherency. This rule dictated that every part of the body of each animal was built by blocks of the same shape. That is to say that all the legs were built by parallelograms or the head by triangles. Thus, we aimed for children to create a learning abstraction during the second task related to their sorting skills. This task followed similar steps:

- The facilitator gave the children one picture at a time and explained the rule of the game. Children as a team tried to replicate the structures on the pictures.
- A worksheet was given to the children to check the characteristics of the last structure.

In the third task (creative construction) the children had to come up with an idea for an animal themselves. They had 5 minutes to built the animal of their choice. All cubes were available to them. However, each child could only have cubes of the same shape. In the end of the activity, the children completed the worksheet and received positive feedback from the facilitator. After this third task, the facilitator thanked the children for participation and took them back to the classroom.

2.4.6 Pilot study

We conducted a pilot-study with children at an after-school care institution to test our research design. In total, three groups of three children participated for about 20 minutes each. In general, children seemed to enjoy the activity. However, we discovered some issues that we needed to address before the actual study:

- Some of the structures that children were asked to build were too complicated for the younger children. It took them much more time than we planned to assemble each structure. As a result, we replaced the challenging structures with simpler ones and we emphasized to children that they always had the choice between a simpler and a more difficult structure.
- Children had to fill in the worksheet that contained some text. This proved to be hard for the young children. As a result, we minimized the amount of written text as much as possible by replacing words with pictures.
- The facilitator was supposed to interfere as little as possible. However, we observed that children tended to seek the facilitator's help each time they faced a challenge on the task. For this reason, we decided to move the facilitator further away from the scene of the task and to allow for interventions only when it was needed for her to give instructions of the procedure as well as during the completion of the worksheets by the children.

2.5 Measures and data analysis

Our analysis was mainly based on the video recordings of children's behaviors. We took into account a total of 103 minutes of video. Unfortunately, due to a technical issue in data recording, we cannot include the third task for group 5 into the analysis. Within the video recordings that were analyzed, we aimed to find behaviors that helped us to address our research questions. Thus, we annotated the videos for relevant behaviors. The annotation scheme and process is described in the following section.

2.5.1 Annotation Scheme

In order to build the annotation scheme, we looked at a subset of the data to identify behaviors that the children displayed in the interaction. We chose this bottom-up approach because we wanted to be open to all possible behaviors that occurred. Our focus was on non-verbal behaviors. When looking at them in the data, we were guided by the initial, very basic, framework derived from relevant literature about children's engagement with a task and their ways of collaboration during activities with shared goals. Three researchers of our team looked at the videos, in order to identify low-level behaviors that indicate children's collaborative activities and engagement in the task.

After a discussion, we agreed that the most promising types of behaviors were children's actions and their spatial position within the playground in relation to the position of the task. In a next step, each of the three researchers compiled a list of actions. After comparison and discussion, we decided to differentiate 21 actions (further explanation in Section 2.5.3).

For the positions of children and task, we tried annotations with different levels of complexity. It turned out that 13 child positions and 7 task positions were feasible and rather reliable (see Figures 2.7 and 2.8, respectively). We annotated the position of each child (C1, C2, and C3) and one position for the task (for further explanation see Section 2.5.4).

In fact, the behaviors that we decided to code based on a first analysis of the data, provide an answer to our first research question (**RQ1**: What kind of non-verbal behaviors appear in children's small-group collaborative activities). The results section will answer how often these behaviors occurred and how they affected collaboration in the groups of children.

Furthermore, we annotated the phases of the task (Task 1 (exploration), 2 (guided construction) and 3 (creative construction)), whether the images in the second task were completed (yes or no), and how many blocks were attached correctly. The last two annotations were only made for the second task, as only in this task the goal was to complete as many constructions as possible in a restricted time. These measures of task performance are further explained in Section 2.5.2.

Together with the coding scheme we developed rules for how to annotate the data to ensure validity and inter-rater reliability:

- All annotations covered the whole duration of an action or the whole time that children / the task remained in one position, respectively.
- The behaviors were coded from the point when they commenced with a movement (e.g., the hand starting to lift up the block to hold it up, the child starting to move to change position).
- The end of the annotation was the start of a new movement.
- Since we decided to consider only children's interactions with each other and not their interaction with the facilitator, only the parts of the sessions when the facilitator was absent were coded.
- The annotations were based on the frontal camera view. If something was not clearly visible from this perspective, one of the other videos was consulted.

The data were annotated by two researchers with the use of ELAN [82]. For the data analysis, we used a Matlab tool developed by the University of Bielefeld to analyze ELAN files. We extracted counts of the children's behaviors for all the sessions and per task.

2.5.2 Task Performance

Next to the children's behaviors, we measured and analyzed the task performance. We wanted to determine how children's collaborative actions and positioning behaviors influenced how well they performed the task. As it has been explained above, task performance was determined only for the second task where the children had a limited time to complete a certain number of images.

Task performance consisted of various measures. We coded in the video files how many constructions the children worked on (1 to 6), whether a certain construction was completed or not (yes/no) according to the given image, the time needed for completion of each task (time in seconds) and the degree of correctness of each structure in relation to the given picture (i.e., percentage of blocks that were in the correct place). As the shapes that the children could assemble had different degrees of difficulty, we also included this measure. More specifically, the degree of difficulty per image was related to the number of blocks that were needed for the image as well as the complexity of the combinations of blocks for the construction. As a result, image 1 had degree of difficulty 1, image 2 had degree of difficulty 2, etc. (see Figure 2.5). Children had the freedom to choose for an image of a variety of degree of difficulty. This is to say that it was not necessary

Block actions	Observation	Attention	Exchange	Aggression
attach block to block; attach block to construction; check block with picture; detach block; looking at block; looking at picture; picking from area; picking from own basket; rotate block	observing other children	showing a block; pointing at a block; pointing at construction; pointing at image	giving a block; receiving a block	grasping from child; grasping from bucket; stopping others from taking blocks

Table 2.2: Action categories

for children to follow any predetermined sequence of pictures. The sum of the degree of difficulty of all images that children replicated indicated the degree of difficulty of the overall task. Since this measure refers to the task, we consider this performance as a collective result of the group of children, rather than as a performance for each child separately.

2.5.3 Actions

As described above, we annotated 21 actions that are explained in Figure 2.6. Although these actions could be seen in isolation, what was more meaningful for our understanding was to examine them in relation to each other. Hence we categorized the actions into 5 categories (see Table 2.2).

Block Actions This category includes actions that are merely task-related. We annotated all the low-level actions that children exhibited during their engagement with the task, such as attaching a block to another.

Observing During the activity, children were not always active in building their construction; instead there were instances when they observed other children's work. We consider these instances important, since, in this

Action	Description
Attachment to block	The child attaches a block to another block, outside the main construction
Attachment to construction	The child attaches an object to the main construction
Check a block with picture	The child checks and compares a block with the picture
Detachment from block	The child detaches a block from another block
Detachment from construction	The child detaches a block or set of block from the main construction
Giving block	A child gives an block to another child either he/she has been asked or not
Grasping from basket	The child grasps a block from another child's basket, without this block being offered
Grasping from child	The child grasps a block from another child, without this block being offered
Looking at the block	The child looks at a block more than 1sec
Looking at the picture	The child looks at the picture
Observing other people's work	The child observe what other children are doing
Picking block from own basket	The child picks a block from his/her own basket
Picking block from the area	The child picks a block from the task area
Pointing block	Child points with the second figure towards a certain block
Pointing construction	The child points with the 2nd finger the construction
Pointing image	Child point with the second finger towards the image
Putting block down	The child puts a block down on the floor
Receiving block	A child receives a block either he/she has asked for it or not
Rotation	The child rotates a block to find a desired position
Showing a block	The child shows a block (with the intention to offer it)
Stopping others from taking a block	The child stops a peer from taking a block

Figure 2.6: Actions that were included in the annotation scheme and the description for each action

age group, children tend to observe adults or more able peers in order to gain insights for their own actions. Hence, we annotated when the children observed others.

Attracting attention These behaviors include children's actions that exhibit their intention for social interaction with other members of the group, e.g., by showing or pointing blocks. These actions can also be interpreted as children's intention to collaborate with other.

Exchange This category includes two actions that relate to pure collaboration that is expressed by only two actions, namely giving a block and receiving a block. These actions exhibit a mutual intention for collaboration from both children. The child who gives the block supports the child who receives the block in building a construction.

Aggression During the task, children sometimes exhibited quite aggressive behaviors towards other members of the group. For example, when children wanted a block from another child, they grasped it, instead of asking for it. These behaviors might cause a defensive behavior from the other child, such as stopping from taking a block.

The above mentioned categories inform us about children's task engagement and their possible types of collaboration with each other. In this research, we defined collaboration as any interaction among the participants that was the result of their mutual intention to achieve a common task-related goal.

In total, 6.233 actions were annotated for the eight groups and three children per session. To analyze the data, we counted the number of occurrences of each action and we calculated the percentage of their duration per action. Additionally, we analyzed the higher-level categories, as they are described in this paragraph (again for number and duration). All of them have been analyzed per session for each child. Additionally, we contextualized the actions and the categories of actions according to the phase of the task, namely exploration, guided construction, and creative construction. Furthermore, we interpreted the results of actions in relation to the task performance for the second phase of the activity (i.e., guided construction) as well as to the position of the children and the position of the task.

2.5.4 Positioning

The preliminary analysis of a subset of data revealed that the positioning of the children in the space is an informative cue of task engagement and children's collaborative behaviors. In fact, the distance between the children and the task and the distance among the children appear to be particularly dynamic in the sessions being an indicator for their task engagement.

Projecting the most frequent positions on a sketch of the playground described in Figure 2.1, thirteen children positions and seven task positions emerged as informative. We observed that those positions together were organized in two circles in the playground: an outer circle comprising the positions distant from the designated position of task in the middle and participants and an inner circle comprising the positions close to the designated position of the task.

The positions were annotated in the data set following a rationale explained in the following paragraphs. The annotation values related to those positions were then analyzed with respect to their counts and duration.

Children positions We defined and annotated the spatial location assumed by the children. During the preliminary analysis of a subset of data, we defined thirteen child positions projecting them into a sketch of the playground (see Figure 2.7).

Each child position was annotated with ELAN following the rules of annotation explained in Section 2.5. The position was coded for each child individually. We considered a position the act of sitting or leaning the upper body toward one of the labeled positions on the playground. To analyze the annotations of the task positions we used SALEM (Statistical AnaLysis of Elan files in Matlab)[31], a Matlab tool to analyze ELAN files developed by the University of Bielefeld. We extracted counts of the child positions for all the sessions and per task. Then, we again provided the results on counts per task.

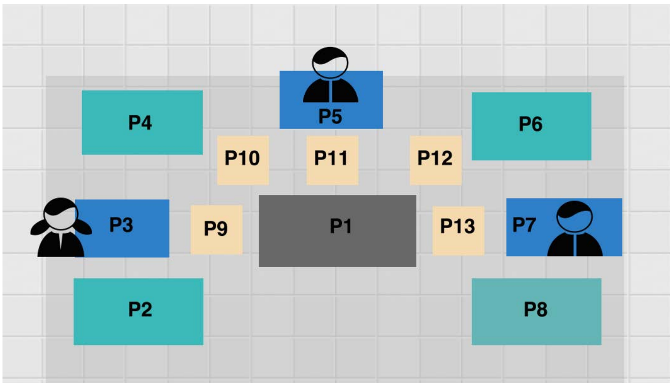


Figure 2.7: Positions of the participants in the playground. P1 represents the designated position of the task. P3, P5, and P7 represent the positions assigned to the participants. P2, P4, P6, and P8 represent the positions that the participants could take in the outer circle. P9, P10, P11, P12, and P13 represent the positions that the participants could take in the inner circle.

Furthermore, we analyzed the number and duration of cases where the position of two or all three children in a group overlapped. This measure informed us about how close the children were to each other, indicating collaboration. Moreover, it provided information about the tendency to cluster in dyads or triads and about possible spatial formations. In order to do so,

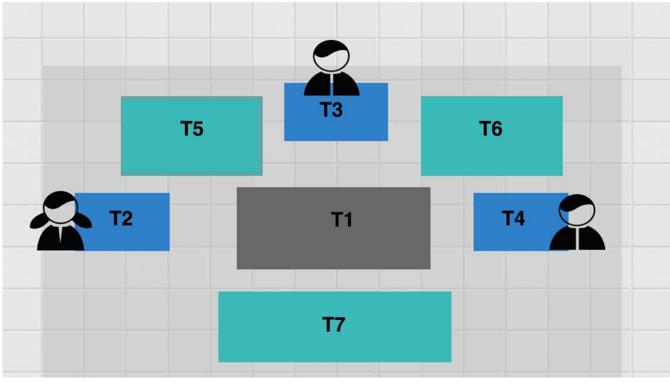


Figure 2.8: Positions of the task in the playground. T1 is the designated position of the task where the facilitator placed the images. T2, T3, and T4 are the designated positions of the participants. T5, T6, and T7 represent possible positions to perform the task outside the designated area.

we defined rules to overlap the task position with multiple child positions. We took either C1/C2/C3 as referent tier and we overlapped it with either one of the other child tier (i.e., C1 referent tier, C2/C3 comparison tier) or two of them (i.e., C1+C2+C3). Following the above rationale, we defined the rules of overlap as follows: overlap between two children in the same position, overlap between the designated position and the opposite positions in the inner circle (e.g., P3+P9 or P3+P10 etc..), overlap between the designated position and the adjacent position in the outer circle (e.g., P3+P4 or P3+P2 etc..), overlap between the inner circle positions and the designated position of the task (i.e., P1). We coded the rules in a script written in Java. The results of the script were loaded in ELAN and the analysis was carried out with the help of SALEM. The results report the counts and percentages of the child positions overlap per task, per group. They are presented by dyads and triads.

Task position As explained above, we also defined and annotated the spatial location where the task was performed (see Figure 2.8). This resulted in one annotation for the whole group. Again, we annotated with ELAN following the rules explained in Section 2.5. To analyze the annotations of the task positions, we again used SALEM. We extracted counts and durations of the task positions annotation for all the sessions and per task. Then, we provided the results on counts for each group. Furthermore, we analyzed the number and duration of the time when the position of the task overlapped with the position of a child. This measure indicated closeness to the task and, hence, the degree of task engagement. We defined rules to overlap the task position with multiple child positions. (i) overlap with designated child position (CP), (ii) overlap with designated position of the task (TP), (iii) overlap with the inner circle position opposite to CP (IC1,IC2, IC3),(iv) overlap with the inner circle position toward the other child (IC12, IC23), (v) overlap with the outer circle position toward the other child (OC12, OC13). We coded rules of overlap in line with the above mentioned categories in a script written in Java. The results of the script were loaded in ELAN and the analysis was carried out with the help of SALEM. The results report the counts and percentages of the overlap per task.

2.6 Results

In the following we present the results of the actions that the children performed, their positioning behavior and the task performance.

2.6.1 Task Performance

Task performance has been considered as a collective result of a group of children, rather than as the performance of each of the children separately. Table 2.3 depicts the number of pictures that each group managed to work on in the time given, the number that they completed successfully, and the percentage of blocks that were put in the right place. We also report the sum of the difficulty of all pictures that the children had worked on (difficulty 1 was assigned to picture 1 and so forth).

The table shows that the performance differed between the groups. The number of shapes completed ranged between 1 and 5. 1 group only completed 1 shape, 4 groups completed 2, 2 groups completed 4, and 1 group 5 shapes. The correct positioning of the blocks ranged between 64% and 100%. Generally this percentage was lower if the group had started to work on an image that they could not finish. The sum of difficulties of pictures that the children had chosen for ranged between 3 (with two images this means that the children chose for the two simplest tasks) and 20. The group with the highest sum of difficulties that performed the best actually consisted of the oldest children (all 8 years old) which is not surprising. Other than that, there was no obvious correlation between performance and age, however, this could also be due to the sample. In the next section, we explain the task performance results some more by putting them in relation to the children's actions.

2.6.2 Actions

In a first step, we analyzed how active each child was in each task. Thereafter, we looked at the categories of actions to learn more about the collaboration between the children. We also related the action results back to the performance results.

group	*number of pictures*	*number of completed pictures*	*% correct blocks*	*sum of difficulty of pictures*
1	2	2	85.00	3
2	4	4	84.00	10
3	3	2	64.00	11
4	5	5	100.00	20
5	2	1	85.50	7
6	3	2	82.00	9
7	3	2	83.30	11
8	4	4	91.25	10

Table 2.3: Task performance in the second task: number of pictures that the children worked on, number that they completed successfully, percentage of blocks put in the correct place, difficulty of the pictures

Activity level per child Table 2.4 depicts the number of actions that each child performed in each task.

Overall, the number of actions that the children performed ranged between 600 and 977 for all three tasks. Hence, there was quite some variation between the groups. However, this difference did not seem to be related to whether the groups were rather balanced or included one child that was much more active than the others. Table 2.4 also shows that overall, there was a tendency that one of the three children performed more actions than the other two. However, in two groups (groups 6 and 7), it seems that two out of three children were more active and one was much less active. In order to find a possible reason for these differences, we checked the age of the children (see Table 2.1). While there is an indication that the older children appear more active, the relatively similar age of the children in the groups does not allow to draw a general conclusion.

It was also noticeable that the activity levels of the children varied with the tasks. In the last task (T3), there was always one child that performed more than 40% of the activities (in just one group it was two because one child was almost inactive). For Task 1, however, this was only true for 4 groups, for Task 2 for 5 groups. Hence, we can assume that the collaboration between the children was different for the creative construction task than for the rest. One child seemed to mostly steer the creative construction, not collaborating as much with the other children as in the other tasks.

While there was this difference between the tasks, in 4 out of 8 groups, one child was most active throughout all tasks (Group 4 - C3 (44%), Group 6 - C1 (41%), Group 7 - C1 (42%) and Group 8 - C1 (42%)). Hence, this child would also be the one leading the creative activity.

Group 7 showed an interesting pattern over time. While there was an increase of the activity of C1 (39% in T1, 40% in T2, 51% in T3), at the same time the activity of C3 decreased (30% in T1, 24% in T2, 14% in T3). This is an indication of the way that one child's dominance may affect the activity of another child.

Another influencing factor might be the nature of the individual tasks. In two groups (group 1 and 6) one child was particularly active on the first and last task (and as a results others were less active) that are more creative in nature and require less collaboration. We will come back to this in the following section when looking at the types of activities that the children performed.

Types of actions In the next step, we analyzed the types of actions to determine which actions the children performed and if there was a relation to their overall level of activity. Table 2.5 depicts the results in terms of number of each action and percentage with respect to all actions. Table 2.6 shows how long each type of action took in relation to the overall duration of all actions.

Overall, all children spent most time on actions related to the blocks. For all children, more than 44% of the actions belonged to this category. The only exception was C2 of Group 4 in the third task, where this child appeared to be almost inactive. In the cases when children seemed highly engaged with the task (i.e., performing many block actions), it was more likely for the other children to observe their work and to try to attract the attention, e.g., by pointing out blocks.

With respect to collaborative actions (i.e., exchanging blocks with others), we found that 2 children did not exhibit such actions at all in any task (Group1/C1 and Group2/C2). Interestingly, 16 out of 24 children did not exhibit any exchange action in at least one of the tasks. In fact, many children seemed to be more likely to perform aggressive actions instead (i.e., to just take blocks from other children or to stop them from doing so). All children displayed an aggressive behavior at some point of the activity.

Children were least likely to exchange blocks in the creative construction task (Task 3) which is in line with the finding that mostly in this task one child was more active than the others. Only for two children the rate of exchanging blocks was highest in this task (C2 and C3 of group 1 obviously exchanged blocks with each other twice and not at all in Task 1 and only once in Task 2). The same was true for the aggressive actions, two children displayed the highest percentage of aggressive actions in Task 3. Again these children belonged to the same group (group 7, C2 with 5 actions (10.64%) and C3 with 3 actions (16.67%)). The children in this group had displayed aggressive behaviors throughout all tasks.

Attracting attention seemed to be most important in the first task. 12 (50%) of the children displayed most of these actions here. We can assume that the children used these actions to show others what they had discovered, e.g., to show them what blocks they had and how they connected to each other. Hence, these actions might have served to establish a common ground between the children.

For the observation actions, there were no clear patterns. Some children seemed to observe more than others but observation actions did not occur more in one task than in the others.

We also want to highlight some findings for individual groups to point out specific group constellations that might appear. In group 1, all children tried to attract the attention for collaboration. In Task 1 this did not result in exchange actions, however, C2 and C3 performed such actions in Task 2 and 3. C1 apparently tried to attract the attention and had an intention for collaboration in all the three tasks but did not succeed. However, potentially as a result, she displayed aggressive behaviors, particularly in Task 2.

In group 3, all three children displayed relatively similar percentages of block related actions. The exception was C3 in Task 2, where she exhibits a higher percentage of block actions (81.29%). In the same task, C1 and C2 exhibit a relatively high percentage of observations and a relatively high percentage of behaviors to attract attention for collaboration. Obviously in the second task, C3 took the lead. This points to the fact that the roles can change for the different tasks.

However, it can also be possible that one person takes the lead throughout the activity as happened in group 6. C1 displayed a high percentage of block-related actions particularly in Tasks 1 and 3, while C2 and C3 observed more. Also the finding that relatively little exchange and aggressive actions took place in this group points to the fact that the children let C1 take a leading role.

Actions and task performance We also want to relate these results back to the performance measures. We focus on the groups that performed best and worst.

There are some things that differentiate group 4 (the best group) from the others. In this group, the children performed more exchange actions than in any other group and task (C1 - 13, C2 - 16, C3 - 16). Also C1 and C2 spend 8% of their time on observing supposedly mostly C 3 who displayed most actions in this task (42%). From these results we can assume that C1 and C2 stayed engaged in the task, observed C3 and helped in handing over missing blocks. This strategy seemed to be successful.

In group 1, the group with the worst performance, all children displayed a similar number of activities. As has been pointed out above already, in this group the children performed quite a lot of aggressive actions (C1 - 15, C2 - 7, C3 - 4) that might have impacted their task performance.

In group 5 (the second-worst group), C1 performed more activities than the others (43%). Also in this group many aggressive behaviors were performed, particularly by C3 (C1 - 2, C2 - 3, C3 - 11). Hence, it seems as if this child did not accept that C1 wanted to take the lead on the task which again might have influenced task performance.

2.6.3 Positioning

To get a further indication of children's task engagement, we analyzed where the children positioned themselves and how close they were to the task.

Child positions The results for the positions of the children are depicted in Table 2.7. In all the tasks, the children tended to remain in the assigned positions (P3 for C1, P5 for C2, and P7 for C3) or to sit/lean toward the inner circle positions opposite to the assigned positions (P9 for C1, P11 for C2, and P13 for C3). In more than 50% of the annotations, the children in each task remained in these positions. There are two exceptions to this: children in position 3 in Task 1 and children in position 1 in Task 2. Children in position 3 in Task 1 also frequently moved to positions P6, P11, and P12, which were the positions closer to C2 in the inner and outer circle. Children in position 1 in Task 2 frequently moved to P1 (the original task position) and P2 (next to the task and away from the other children).

Table 2.8 depicts the number of cases and percentages when the position of two or all three children overlapped, indicating how often the children shared the same position and supposedly worked together. In most cases we found that only two children shared the same position and, thus, worked together only in dyads. This is in line with the findings on the activity levels.

We want to highlight the results for group 3 because it was really different in that all children were in the same space in the first task, in the second task the overlap is between C1 and C2, and C2 and C3, and in the third task there are no overlaps between the children at all. Hence, it seems as if the children became dis-engaged with each other over time or stayed further away from each other because of the nature of the task.

Also group 6 stood out in that the overlap occurred to a very high percentage between C1 and C2 throughout the tasks (T1: 71%, T2: 89%, T3: 79%). Hence this measure underlines the finding that these two children were most active.

Child positions and task performance In group 1, that performed worst in Task 2, in Task 1 and 3, C2 shared the space with C1 (overlap 34% and 42%, respectively) but also with C3 to a similar extent (39% and 37%). In Task 2, C1 mainly stayed close to C2 (50%), but C2 was not so close to

C3 (9%). An assumption could be that the lack of overlap between these two children and a change in collaboration compared to the other tasks was one cause for the weak performance of the group.

Group 5 had also performed below average. However, in this case there was quite some overlap between C1 and C3 (32%) and between C2 and C3 (32%). Hence, all the children seemed to be close to at least C1, who, as we discussed above, was most active. However, as we also mentioned, there was a lot of aggressive behavior in this group such that the closeness alone cannot indicate good collaboration.

In fact, the percentage of overlap between the children was not much higher for the best group (group 4) with 37% between C1 and C2 and 40% between C2 and C3. However, as it has been explained above, the actions that took place between these children were less aggressive and more collaborative.

Task position For the task position, we first had a look at Table 2.9. Generally we observed that the task was mostly located in T1, hence, in the designated position of the task for Tasks 2 and 3 (33.46% and 41.82% of the time, respectively). However, in the first task, the constructions were performed more often in the designated positions of one of the children (22.56% at T1 (which is next to C1), 28.86% in T2 (which is the original position of C1), 10.42% in T3 (which is the original position of C2), and 25.86% in T4 (which is the original position of C3)). The task was barely performed in T5 or T6 which are the positions between C1 and C2, C2 and C3, respectively. It was never performed in T7 which was the position furthest away from the children. These results may indicate that children explored the objects rather by themselves than in a shared space; whereas they chose the central, shared space for more collaborative tasks.

One of the most outstanding groups again was group 6 (see Table 2.10). In this group, the task position barely changed. It stayed in the same place in Tasks 1 and 3 (see Tables 2.11 and 2.13) and only switched to another location and back in Task 2 (see Table 2.12). The task was mostly performed in T2 which is closest to the original position of C1. When the task switched location in Task 2 it went to position T5 (40% of the overall duration) which is between C1 and C2. This is in line with the fact that C1 was most active in the group as has been explained above.

	nr all actions	*nr actions T1*	*nr actions T2*	*nr actions T3*
Group 1				
C1	181 (.30)	65 (.35)	80 (.36)	36 (.19)
C2	259 (.43)	82 (.44)	74 (.33)	103 (.54)
C3	160 (.27)	38 (.21)	70 (.31)	52 (.27)
all	600 (1.00)	185 (1.00)	224 (1.00)	191 (1.00)
Group 2				
C1	258 (.32)	51 (.27)	138 (.37)	69 (.29)
C2	286 (.36)	70 (.37)	115 (.31)	101 (.42)
C3	259 (.32)	69 (.36)	119 (.32)	71 (.29)
all	803 (1.00)	190 (1.0)	372 (1.0)	241 (1.0)
Group 3				
C1	235 (.31)	89 (.36)	93 (.26)	53 (.34)
C2	222 (.29)	83 (.33)	110 (.31)	29 (.19)
C3	306 (.40)	78 (.31)	155 (.43)	73 (.47)
all	763 (1.00)	250 (1.0)	358 (1.0)	155 (1.0)
Group 4				
C1	249 (.28)	76 (.27)	132 (.26)	41 (.44)
C2	252 (.28)	89 (.31)	162 (.32)	1 (.01)
C3	388 (.44)	121 (.42)	216 (.42)	51 (.55)
all	889 (1.0)	286 (1.0)	510 (1.0)	93 (1.0)
Group 5				
C1	327 (.41)	112 (.36)	215 (.43)	0 (0)
C2	201 (.25)	79 (.26)	122 (.24)	0 (0)
C3	279 (.35)	117 (.38)	162 (.32)	0 (0)
all	807 (1.0)	308 (1.0)	499 (1.0)	0 (0)
Group 6				
C1	396 (.41))	146 (.46)	160 (.35)	90 (.44)
C2	346 (.35)	127 (.40)	152 (.34)	67 (.33)
C3	235 (.24)	46 (.14)	140 (.31)	49 (.24)
all	977 (1.0)	319 (1.0)	452 (1.0)	206 (1.0)
Group 7				
C1	357 (.42)	79 (.39)	210 (.40)	68 (.51)
C2	295 (.34)	63 (.31)	185 (.36)	47 (.35)
C3	205 (.24)	62 (.30)	125 (.24)	18 (.14)
all	857 (1.0)	204 (1.0)	520 (1.0)	133 (1.0)
Group 8				
C1	280 (.42)	105 (.40)	141 (.45)	34 (.56)
C2	178 (.28)	76 (.29)	81 (.28)	21 (.35)
C3	182 (.30)	84 (.31)	92 (.28)	6 (.09)
all	640 (1.00)	265 (1.00)	314 (1.00)	61 (1.00)

Table 2.4: Number (percentages) of actions per child, per child in each task, and per group.

		block actions	*exchange*	*aggression*	*attention*	*observation*
Group 1						
C1	T1	49 (75.38)	0 (0)	4 (6.15)	3 (4.62)	9 (13.85)
	T2	51 (63.75)	0 (0)	15 (18.75)	6 (7.50)	8 (10.00)
	T3	24 (66.67)	0 (0)	1 (2.78)	5 (13.89)	6 (16.67)
C2	T1	64 (78.05)	0 (0)	5 (6.10)	6 (7.32)	7 (8.54)
	T2	56 (75.68)	1 (1.35)	7 (9.46)	4 (5.41)	6 (8.11)
	T3	93 (90.29)	2 (1.94)	4 (3.88)	2 (1.94)	2 (1.94)
C3	T1	22 (57.89)	0 (0)	1 (2.63)	4 (10.53)	11 (28.95)
	T2	36 (51.43)	1 (1.43)	4 (5.71)	11 (15.71)	18 (25.71)
	T3	36 (69.23)	2 (3.85)	0 (0)	5 (9.62)	9 (17.31)
Group 2						
C1	T1	39 (76.47)	0 (0)	7 (13.73)	2 (3.92)	3 (5.88)
	T2	111 (80.43)	3 (2.17)	14 (10.14)	4 (2.90)	6 (4.35)
	T3	57 (82.61)	0 (0)	0 (0)	3 (4.35)	9 (13.04)
C2	T1	58 (82.86)	0 (0)	6 (8.57)	0 (0)	6 (8.57)
	T2	81 (70.43)	0 (0)	5 (4.35)	3 (2.61)	26 (22.61)
	T3	95 (94.06)	0 (0)	1 (0.99)	0 (0)	5 (4.95)
C3	T1	46 (66.67)	1 (1.45)	5 (7.25)	1 (1.45)	16 (23.19)
	T2	77 (64.71)	4 (3.36)	9 (7.56)	2 (1.68)	27 (22.69)
	T3	51 (71.83)	0 (0.00)	3 (4.23)	2 (2.82)	15 (21.13)
Group 3						
C1	T1	60 (67.42)	4 (4.49)	11 (12.36)	6 (6.74)	8 (8.99)
	T2	41 (44.09)	7 (7.53)	3 (3.23)	22 (23.66)	20 (21.51)
	T3	34 (64.15)	3 (5.66)	0 (0)	6 (11.32)	10 (18.87)
C2	T1	51 (61.45)	1 (1.20)	8 (9.64)	5 (6.02)	18 (21.69)
	T2	73 (66.36)	3 (2.73)	7 (6.36)	11 (10.00)	16 (14.55)
	T3	16 (55.17)	1 (3.45)	1 (3.45)	3 (10.34)	8 (27.59)
C3	T1	44 (56.41)	6 (7.69)	13 (16.67)	9 (11.54)	6 (7.69)
	T2	126 (81.29)	6 (3.87)	14 (9.03)	3 (1.94)	6 (3.87)
	T3	58 (79.45)	3 (4.11)	3 (4.11)	4 (5.48)	5 (6.85)
Group 4						
C1	T1	41 (53.95)	8 (10.53)	0 (0)	6 (7.89)	21 (27.63)
	T2	78 (59.09)	13 (9.85)	3 (2.27)	4 (3.03)	34 (25.76)
	T3	36 (87.80)	0 (0)	0 (0	0 (0)	5 (12.20)
C2	T1	49 (55.06)	4 (4.49)	2 (2.25)	3 (3.37)	31 (34.83)
	T2	85 (52.47)	16 (9.88)	7 (4.32)	19 (11.73)	35 (21.60)
	T3	0 (0)	0 (0)	0 (0)	0 (0)	1 (100)
C3	T1	111 (91.74)	5 (4.13)	4 (3.31)	1 (0.83)	0 (0)
	T2	190 (87.96)	16 (7.41)	9 (4.17)	1 (0.46)	0 (0)
	T3	47 (92.16)	0 (0)	1 (1.96)	0 (0)	3 (5.88)

		block actions	*exchange*	*aggression*	*attention*	*observation*
		Group 5				
C1	T1	83 (74.11)	7 (6.25)	1 (0.89)	11 (9.82)	10 (8.93)
	T2	174 (80.93)	7 (3.26)	2 (0.93)	11 (5.12)	21 (9.77)
	T3	-	-	-	-	-
C2	T1	60 (75.95)	1 (1.27)	1 (1.27)	4 (5.06)	13 (16.46)
	T2	98 (80.33)	2 (1.64)	3 (2.46)	1 (0.82)	18 (14.75)
	T3	-	-	-	-	-
C3	T1	71 (60.68)	5 (4.27)	2 (1.71)	16 (13.68)	23 (19.66)
	T2	115 (70.99)	9 (5.56)	11 (6.79)	9 (5.56)	18 (11.11)
	T3	-	-	-	-	-
		Group 6				
C1	T1	134 (91.78)	3 (2.05)	4 (2.74)	2 (1.37)	3 (2.05)
	T2	148 (92.50)	0 (0)	6 (3.75)	2 (1.25)	4 (2.50)
	T3	88 (97.78)	0 (0)	1 (1.11)	0 (0)	1 (1.11)
C2	T1	111 (87.40)	1 (0.79)	7 (5.51)	6 (4.72)	2 (1.57)
	T2	141 (92.76)	0 (0)	2 (1.32)	1 (0.66)	8 (5.26)
	T3	64 (95.52)	0 (0)	0 (0)	0 (0)	3 (4.48)
C3	T1	31 (67.39)	3 (6.52)	1 (2.17)	3 (6.52)	8 (17.39)
	T2	133 (95.00)	0 (0)	2 (1.43)	2 (1.43)	3 (2.14)
	T3	41 (83.67)	0 (0)	1 (2.04)	0 (0)	7 (14.29)
		Group 7				
C1	T1	54 (68.35)	3 (3.80)	4 (5.06)	12 (15.19)	6 (7.59)
	T2	150 (71.43)	7 (3.33)	22 (10.48)	15 (7.14)	16 (7.62)
	T3	54 (79.41)	1 (1.47)	4 (5.88)	6 (8.82)	3 (4.41)
C2	T1	40 (63.49)	2 (3.17)	4 (6.35)	5 (7.94)	12 (19.05)
	T2	135 (72.97)	11 (5.95)	19 (10.27)	10 (5.41)	10 (5.41)
	T3	35 (74.47)	0 (0)	5 (10.64)	6 (12.77)	1 (2.13)
C3	T1	46 (74.19)	3 (4.84)	3 (4.84)	4 (6.45)	6 (9.68)
	T2	88 (70.40)	6 (4.80)	9 (7.20)	5 (4.00)	17 (13.60)
	T3	10 (55.56)	1 (5.56)	3 (16.67)	3 (16.67)	1 (5.56)
		Group 8				
C1	T1	90 (85.57)	4 (3.80)	8 (7.76)	1 (0.95)	2 (1.90)
	T2	130 (92.19)	4 (2.84)	6 (4.25)	0 (0)	1 (0.70)
	T3	31 (92.18)	0 (0)	0 (0)	0 (0)	2 (5.88)
C2	T1	54 (71.05)	0 (0)	3 (3.95)	5 (6.58)	14 (18.42)
	T2	45 (55.55)	6 (7.41)	4 (9.94)	8 (19.51)	18 (22.22)
	T3	10 (47.62)	1 (4.76)	1 (4.76)	3 (14.29)	6 (28.57)
C3	T1	53 (63.09)	5 (5.95)	6 (7.14)	8 (9.52)	12 (14.29)
	T2	59 (64.13)	4 (4.35)	5 (5.43)	14 (15.22)	10 (10.87)
	T3	3 (16.67)	0 (0)	1 (16.67)	1 (16.67)	3 (50.00)

Table 2.5: Number of actions (percentage of all actions) that the children spent on tasks belonging to the types of collaboration.

		block actions	*exchange*	*aggression*	*attention*	*observation*
Group 1						
C1	T1	18.92	0.00	1.54	1.16	3.47
	T2	13.42	0.00	3.95	1.58	2.11
	T3	10.96	0.00	0.46	2.28	2.74
C2	T1	24.71	0.00	1.93	2.32	2.70
	T2	14.74	0.26	1.84	1.05	1.58
	T3	42.47	0.91	1.83	0.91	0.91
C3	T1	8.49	0.00	0.39	1.54	4.25
	T2	9.47	0.26	1.05	2.89	4.74
	T3	16.44	0.91	0.00	2.28	4.11
Group 2						
C1	T1	24.38	0.00	4.38	1.25	1.88
	T2	28.76	0.78	3.63	1.04	1.55
	T3	24.15	0.00	0.00	1.27	3.81
C2	T1	36.25	0.00	3.75	0.00	3.75
	T2	20.98	0.00	1.30	0.78	6.74
	T3	40.24	0.00	0.42	0.00	2.12
C3	T1	28.75	0.63	3.13	0.63	10.00
	T2	19.95	1.04	2.33	0.52	6.99
	T3	21.61	0.00	1.27	0.85	6.36
Group 3						
C1	T1	23.53	1.57	4.31	2.35	3.14
	T2	11.71	2.00	0.86	6.29	5.71
	T3	20.48	1.81	0.00	3.61	6.02
C2	T1	20.00	0.39	3.14	1.96	7.06
	T2	20.86	0.86	2.00	3.14	4.57
	T3	9.64	0.60	0.60	1.81	4.82
C3	T1	17.25	2.35	5.10	3.53	2.35
	T2	36.00	1.71	4.00	0.86	1.71
	T3	34.94	1.81	1.81	2.41	3.01
Group 4						
C1	T1	16.27	3.17	0.00	2.38	8.33
	T2	17.81	2.97	0.68	0.91	7.76
	T3	31.58	0.00	0.00	0.00	4.39
C2	T1	19.44	1.59	0.79	1.19	12.30
	T2	19.41	3.65	1.60	4.34	7.99
	T3	0.00	0.00	0.00	0.00	0.88
C3	T1	44.05	1.98	1.59	0.40	0.00
	T2	43.38	3.65	2.05	0.23	0.00
	T3	41.23	0.00	0.88	0.00	2.63

		block actions	*exchange*	*aggression*	*attention*	*observation*
Group 5						
C1	T1	29.75	2.51	0.36	3.94	3.58
	T2	56.68	2.28	0.65	3.58	6.84
	T3	-	-	-	-	-
C2	T1	21.51	0.36	0.36	1.43	4.66
	T2	31.92	0.65	0.98	0.33	5.86
	T3	-	-	-	-	-
C3	T1	25.45	1.79	0.72	5.73	8.24
	T2	37.46	2.93	3.58	2.93	5.86
	T3	-	-	-	-	-
Group 6						
C1	T1	59.03	1.32	1.76	0.88	1.32
	T2	62.71	0.00	2.54	0.85	1.69
	T3	80.00	0.00	0.91	0.00	0.91
C2	T1	48.90	0.44	3.08	2.64	0.88
	T2	59.75	0.00	0.85	0.42	3.39
	T3	58.18	0.00	0.00	0.00	2.73
C3	T1	13.66	1.32	0.44	1.32	3.52
	T2	56.36	0.00	0.85	0.85	1.27
	T3	37.27	0.00	0.91	0.00	6.36
Group 7						
C1	T1	34.18	1.90	2.53	7.59	3.80
	T2	29.30	1.37	4.30	2.93	3.13
	T3	27.64	0.51	2.05	3.08	1.54
C2	T1	25.32	1.27	2.53	3.16	7.59
	T2	26.37	2.15	3.71	1.95	1.95
	T3	17.95	0.00	2.56	3.08	0.51
C3	T1	29.11	1.90	1.90	2.53	3.80
	T2	17.19	1.17	1.76	0.98	3.32
	T3	5.13	1.51	1,54	1.54	0.51
Group 8						
C1	T1	30.51	1.36	2.71	0.34	0.68
	T2	35.42	1.09	1.63	0.00	0.27
	T3	40.79	0.00	0.00	0.00	2.63
C2	T1	18.31	0.00	1.02	1.69	4.75
	T2	12.26	1.63	1.09	2.18	4.90
	T3	13.16	1.32	1.32	3.95	7.89
C3	T1	17.97	1.69	2.03	2.71	4.07
	T2	16.08	1.09	1.36	3.81	2.72
	T3	3.95	0.00	1.32	1.32	3.95

Table 2.6: Rate of duration of actions of overall duration of all actions.

Child Position									
Task		*1*			*2*			*3*	
	C1(%)	**C2(%)**	**C3(%)**	**C1(%)**	**C2(%)**	**C3(%)**	**C1(%)**	**C2(%)**	**C3(%)**
P1	31 (16.40)	10 (5.32)	23 (8.85)	81 (20.82)	21 (7.05)	54 (13.24)	26 (18.18)	0	19 (14.62)
P2	4 (2.12)	0	0	61 (15.68)	1 (0.34)	1 (.25)	2 (1.49)	0	0
P3	52 (27.51)	1 (0.53)	0	37 (9.51)	1 (0.34)	1 (0.25)	46 (32.17)	3 (4.05)	0
P4	5 (2.65)	16 (8.51)	2 (0.77)	32 (8.23)	27 (9.06)	1(0.25)	6 (4.20)	7 (9.46)	1 (0.77)
P5	2 (1.06)	46 (24.47)	1 (0.38)	5 (1.29)	77 (25.84)	2 (0.49)	1 (0.70)	21 (28.38)	0
P6	1 (0.53)	39 (20.74)	37 (14.23)	0	4 (1.34)	38 (9.31)	0	6 (8.11)	8 (6.15)
P7	0	0	41 (15.77)	0	2 (0.67)	97 (23.77)	1 (0.70)	0	37 (28.46)
P8	0	0	0	0	0	0	0	0	2 (1.54)
P9	59 (31.22)	2 (1.06)	4 (1.54)	96 (24.68)	7 (2.35)	4 (0.98	37 (25.87)	3 (4.05)	0
P10	18 (9.52)	12 (6.38)	6 (2.31)	42 (10.80)	40 (13.42)	6 (1.47)	14 (9.79)	7 (9.46)	4 (3.08)
P11	13 (6.88)	44 (23.40)	51 (19.62)	22 (5.66)	78 (26.17)	32 (7.84)	5 (3.50)	19 (25.68)	8 (6.15)
P12	0	13 (6.91)	44 (16.92)	0	34 (11.41)	65 (15.93)	0	8 (10.81)	15 (11.54)
P13	4 (2.12)	5 (2.66)	51 (19.62)	13 (3.34)	6 (2.01)	107 (26.23)	5 (3.50)	0	36 (27.69)
	189 (100)	188 (100)	260 (100)	389 (100)	298 (100)	408 (100)	143 (100)	74 (100)	130 (100)
all		637			1.095			347	

Table 2.7: Counts and percentages of the child positions per task.

Task position and child position Table 2.14 shows the percentages of the overlap of child and task positions, in other words the time when the children were close to the task. Comparing this information with the level of activity of each child, we find a strong connection between being more active and being closer to the task for a longer time. Hence, the positioning information supports the information on the level of activity. We want to illustrate this with the example of group 6. The task position for this group mainly overlaps with the position of C1 who was the most active child but also with C2. What we could observe most frequently was that C2 was leaning toward him in the outer circle and C3 was leaning toward him through the designated position of the task.

Position overlap and task performance We again look at the groups that performed best and worst.

For group 4, performing best, the task position overlapped mainly with the position of C3 who was most active in that task (56%) (see Table 2.14). Overall, the task stayed in position T4, with C3, for 77% of the time. We observed that the other children mostly leaned towards this child and this position.

For group 1, performing worst, the task mainly stayed in T1 (in the middle, 36% of the time) and T3 (with C2, 30%). Hence, the task position mainly overlapped with the position of C2 (68% of the annotations). Again, we observed that the other children often leaned toward this child being close to the task (C1 - 32% and C3 - 45% of the annotations). Hence, all children were engaged in the task, but the collaboration was still not going well as has been mentioned above.

For group 5, that also performed below average, the task position was located mostly in T2 (with C1, 43% of the time) and T3 (with C2, 40% of the time). Interestingly these children did not choose the position between them. However, the annotations of all children's positions, including C3, overlap with the task space to a high degree (C1 - 70%, C2 - 43%, C3 - 54%). So again, all children seemed rather engaged. However, maybe the fact that the task was always performed in the location of one of the children, rather than in a shared space, points to the conclusion that the collaboration was not going so well in this group.

	dyad/ triad	*Task 1 nr. overlap (%)*	*Task 2 nr. overlap (%)*	*Task 3 nr. overlap (%)*
Group 1	c1/c2	**15 (34.09)**	**37 (50)**	**8 (42.10)**
	c2/c3	**17 (38.63)**	7 (9.45)	**7 (36.84)**
	c1/c3	5 (11.36)	20 (27.02)	3 (15.78)
	c1/c2/c3	7 (15.90)	10 (13.51)	1 (5.26)
	sum	44 (100)	74 (100)	19 (100)
Group 2	c1/c2	**57 (64.04)**	**92 (28.75)**	**24 (60)**
	c2/c3	7 (7.86)	54 (16.87)	7 (17.5)
	c1/c3	5 (5.61)	**107 (33.43)**	6 (15)
	c1/c2/c3	20 (22.47)	67 (20.93)	3
	sum	89 (100)	320 (100)	40 (100)
Group 3	c1/c2	30 (24.39)	**72 (29.62)**	0
	c2/c3	32 (26.01)	**73 (30)**	0
	c1/c3	28 (22.76)	51 (20.98)	0
	c1/c2/c3	33 (26.82)	47 (19.34)	0
	sum	123 (100)	243 (100)	0
Group 4	c1/c2	**18 (42.85)**	**34 (36.95)**	0
	c2/c3	9 (21.42)	**37 (40.21)**	**2 (40)**
	c1/c3	10 (23.80)	8 (8.69)	**3 (60)**
	c1/c2/c3	5 (11.90)	13 (14.13)	0
	sum	42 (100)	92 (100)	5 (100)
Group 5	c1/c2	25 (22.12)	23 (19.65)	-
	c2/c3	**42 (37.16)**	**37 (31.62)**	-
	c1/c3	**32 (28.13)**	**38 (32.47)**	-
	c1/c2/c3	14 (12.38)	19 (16.23)	-
	sum	113 (100)	117 (100)	-
Group 6	c1/c2	**25 (71.42)**	**24 (88.88)**	**11 (78.57)**
	c2/c3	4 (11.42)	1 (3.70)	1 (7.1)
	c1/c3	2 (5.71)	1 (3.70)	2 (14.28)
	c1/c2/c3	4 (11.42)	1 (3.70)	0
	sum	35 (100)	27 (100)	14 (100)
Group 7	c1/c2	**8 (50)**	**39 (97.5)**	**20 (41.66)**
	c2/c3	4 (25)	0	**21 (43.75)**
	c1/c3	2 (12.5)	1 (2.5)	7 (14.58)
	c1/c2/c3	2 (12.5)	0	0
	sum	16 (100)	40 (100)	48 (100)
Group 8	c1/c2	**16 (30.18)**	**53 (44.53)**	5 (38.46)
	c2/c3	**20 (37.7)**	22 (18.48)	0
	c1/c3	13 (24.52)	19 (15.96)	**8 (51.53)**
	c1/c2/c3	4 (7.54)	25 (21)	0
	sum	53 (100)	119 (100)	13 (100)

Table 2.8: Number and percentage of overlap between dyads (c1/c2, c2/c3, c1/c3) and triads (c1/c2/c3) per group per task.

Task Position						
1st task: Exploration						
Position	*Count*	*%*	*Dur mean*	*SD*	*Dur sum*	*%*
T1	16	6.63	21.10	36.54	337.6	22.56
T2	18	7.46	23.97	26.44	432.0	28.86
T3	15	6.22	10.37	12.46	156.0	10.42
T4	18	7.46	21.54	29.86	387.0	25.86
T5	7	2.90	12.34	12.86	86.1	5.75
T6	10	4.14	9.82	5.69	98.0	6.55
T7	0	0	0	0	0	0
2nd task Guided Construction						
T1	40	16.60	24.32	37.78	972.0	33.46
T2	11	4.56	39.72	51.13	436.7	15.03
T3	24	9.95	29.44	45.14	705.6	24.29
T4	24	9.95	17.13	17.33	410.4	14.13
T5	11	4.53	26.24	31.09	288.2	9.92
T6	14	5.80	6.56	7.78	92.4	3.18
T7	0	0	0	0	0	0
3rd task Creative Construction						
T1	20	8.29	28.29	26.23	565.8	41.82
T2	13	5.39	32.88	36.00	427.7	31.61
T3	10	4.14	12.29	13.42	123.0	9.09
T4	10	4.14	9.55	6.77	96.0	7.09
T5	1	0.41	84.82	0	84.8	6.27
T6	3	1.24	18.65	31.25	55.8	4.12
T7	0	0	0	0	0	0

Table 2.9: Counts and durations of task positions in all the sessions. The results are provided per task.

Task Position						
	Group 1			*Group 2*		
	Task1(%)	Task 2 (%)	Task 3(%)	Task1(%)	Task 2(%)	Task 3(%)
T1	5 (31.25)	6 (30)	8 (53.33)	2(40)	6 (50)	5 (38.46)
T2	1(6,25)	2 (10)	(13.33)	0 (0)	0 (0)	2 (15,38)
T3	6 (37.5)	6(30)	1 (6,66)	1(20)	3 (25)	3 (23.07)
T4	3(18.75)	4 (20)	3(15)	2(40)	0 (0)	2(15.38)
T5	0 (0)	2 (10)	0 (0)	0 (0)	2 (16.66)	0 (0)
T6	1 (6.25)	0 (0)	1(6.66)	0 (0)	1(8.33)	1 (7.69)
all.	16 (100)	20 (100)	15 (100)	5 (100)	12 (100)	13 (100)
	Group 3			*Group 4*		
	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
T1	3 (11.53)	7 (43.75)	2 (50)	0 (0)	10 (29.41)	1 (33.33)
T2	9 (34.61)	0 (0)	0 (0)	0 (0)	1 (2.94)	0 (0)
T3	2 (7.69)	3 (18.75)	1 (25)	0 (0)	0 (0)	0 (0)
T4	2 (7.69)	0 (0)	1(25)	4 (57.14)	16 (47.05)	2 (66.66)
T5	4 (15.38)	2 (12.5)	0 (0)	0 (0)	0 (00	0 (0)
T6	6 (23.07)	4 (25)	0 (0)	3 (42.85)	7 (20.58)	0 (0)
all.	26 (100)	16 (100)	4 (100)	7 (100)	34 (100)	3 (100)
	Group 5			*Group 6*		
	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
T1	3 (17.64)	3 (43.85)	0	0 (0)	0 (0)	0 (0)
T2	4 (23.52)	2 (28.57)	0	1 (100)	2 (66.66)	1 (100)
T3	4 (23.52)	1 (14.28)	0	0 (0)	0 (0)	0 (0)
T4	6(35.29)	1(14.28)	0	0 (0)	0 (0)	0 (0)
T5	0 (0)	0 (0)	0	0 (0)	1 (33.33)	0 (0)
T6	0 (0)	0 (0)	0	0 (0)	0 (0)	0 (0)
all	17(100)	7(100)	0	1(100)	3 (100)	1(100)
	Group 7			Group 8		
	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
T1	3 (30)	2 (16.66)	0 (0)	3 (20)	7 (46.66)	1 (50)
T2	0 (0)	0 (0)	0 (0)	7 (46.66)	2 (13.33)	0 (0)
T3	3 (30)	5 (41.66)	1 (33.33)	4 (26.66)	4 (26.66)	0 (0)
T4	1 (10)	1	0 (0)	1(6.66)	2(13.33)	1 (50)
T5	3 (30)	4 (33.33)	1 (33.33)	0 (0)	0 (0)	0 (0)
T6	0 (0)	0 (0)	1(33.33)	0 (0)	0 (0)	0 (0)
T6	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
all	10 (100)	12 (100)	3 (100)	15 (100)	15 (100)	2 (100)

Table 2.10: Counts and percentages of task positions. The results are presented per group and per task.

Task Position (durations)						
	Group 1			*Group 2*		
	Dur sum/%	*M(s)*	*SD(s)*	*Dur sum /%*	*M (s)*	*SD (s)*
T1	133.55/61.13	26.71	20.69	160.6/94.88	80.30	93.43
T2	9.70/4.44	9.70	0	0	0	0
T3	43.68/20	7.28	5.25	1.50/0.88	1.50	0
T4	21.84/9.99	17.10	9.45	7.16/4.24	3.58	1.63
T5	0	0	0	0	0	0
T6	9.68/4.44	9.68	0	0	0	0
	Group 3			*Group 4*		
	Dur sum/%	M(s)	SD(s)	Dur sum/%	M (s)	SD (s)
T1	17.91/7.28	5.97	2.12	0	0	0
T2	102.15/41.39	11. 35	8.18	0	0	0
T3	10.88/4.4	5.44	3.05	0	0	0
T4	22.42/9.04	11.21	4.17	206.28/84	51.57	41.77
T5	44.24/17.92	11.06	17.20	0	0	0
T6	49.14/19.97	8.19	5.98	39.3/16	13.10	5.66
	Group 5			*Group 6*		
	Dur sum/ %	M (s)	SD (s)	Dur sum/%	M (s)	SD (s)
T1	12.51/5.5	4.17	2.51	0	0	0
T2	101.48/45	25.37	33.28	64.13/100	64.13	0
T3	45.28/20	11.32	6.14	0	0	0
T4	65.76/29.5	10.96	17.58	0	0	0
T5	0	0	0	0	0	0
T6	0	0	0	0	0	0
	Group 7			Group 8		
	Dur sum/%	M (s)	SD (s)	Dur sum/%	M (s)	SD (s)
T1	12.99/7.56	4.33	23.22/8.48	2.4	7.74	1.24
T2	0	0	0	224/81.82	32.00	32.67
T3	81.83/47.7	27.11	34.12	21.88/8	5.47	6.11
T4	34.74/20.24	34.74	0	4.65/1.7	4.65	0
T5	42.06/24.5	14.02	6.67	0	0	0
T6	0	0	0	0	0	0

Table 2.11: Durations of task positions in **task 1**. Overall duration, mean and standard deviation are expressed in seconds.

Task Position (durations)						
	Group 1			*Group 2*		
	Dur sum/ %	*M(s)*	*sd(s)*	*Dur sum / %*	*M (s)*	*sd (s)*
T1	112.74/35.55	18.79	8.07	360.72/90.63	60.12	62.68
T2	41.86/13.59	20.93	7.30	0	0	0
T3	93.48/30.3	15.58	8.77	30.03/7.54	10.01	5.61
T4	41.2/13.35	10.03	6.33	0	0	0
T5	19.16/6.21	9.80	9.74	5.3/1.33	2.65	0.037
T6	0	0	0	2.02/0.5	2.02	0
	Group 3			*Group 4*		
	Dur sum / %	M(s)	sd(s)	Dur sum / %	M (s)	sd (s)
T1	317.73/86.6	45.39	52.64	45.6/10.48	4.56	4.54
T2	0	0	0	6.79/1.57	6.79	0
T3	16.08/4.38	5.36	3.04	0	0	0
T4	0	0	0	335,2/77,17	20.95	20.25
T5	5.7/1.55	2.85	0.70	0	0	0
T6	28.25/7.67	7.10	2.35	46.83/0	6.69	8.08
	Group 5			*Group 6*		
	Dur sum / %	M (s)	sd (s)	Dur sum / %	M (s)	sd (s)
T1	64/15.4	21.6	28.09	0	0	0
T2	179.72/43.26	89.86	123.67	66.02/67.76	66.89	1.23
T3	167.39/40.29	167.39	0	0	0	0
T4	4.37/1.05	4.37	0	0	0	0
T5	0	0	0	87.92/39.65	87.92	0
T6	3.67/24.38	14.02	14.02	0	0	0
	Group 7			Group 8		
	Dur sum / %	M (s)	sd (s)	Dur sum/%	M (s)	sd (s)
T1	26.94/6.38	13.47	14.58	51.45/19.50	7.35	6.51
T2	0	0	0	47.88/18.12	23.94	14.11
T3	207.4/49.12	41.48	56.22	144.52/54.72	36.13	43.88
T4	17.76/4.2	17.76	0	20.25/7.66	10.25	9.68
T5	170.12/40.3	42.53	25.53	0	0	0
T6	0	0	0	0	0	0
T6	0	0	0	0	0	0

Table 2.12: Durations of task positions in **task 2**. Overall duration, mean and standard deviation are expressed in seconds.

Task Position (durations)						
	Group 1			*Group 2*		
	Dur sum/%	*M(s)*	*sd(s)*	*Dur sum/%*	*M (s)*	*sd (s)*
T1	118/59.78	14.75	7.38	127.4/54.19	25.48	31.08
T2	18.02/9.12	9.10	0.77	42.44/18.04	21.22	24.77
T3	10.25/5.2	10.25	0	42.51/18.07	14.17	9.37
T4	32.49/16.45	10.83	5.26	17.7/7.53	8.85	5.00
T5	0	0	0	0	0	0
T6	18.64/9.45	18.64	0	5.09/2.17	5.09	0
	Group 3			*Group 4*		
Dur sum/ %		M(s)	sd(s)	Dur sum/%	M (s)	sd (s)
T1	138.06/83.22	69.03	17.26	48.7/80.33	48.70	28.33
T2	0	0	0	0	0	0
T3	3.50/2.1	3.50	0	0	0	0
T4	24.32/14.68	24.32	0	11.92/19.67	5.96	9.67
T5	0	0	0	0	0	0
T6	0	0	0	0	0	0
	Group 5			*Group 6*		
	Dur sum/%	M (s)	sd (s)	Dur sum/%	M (s)	sd (s)
T1	-	-	-	0	0	0
T2	-	-	-	110.70/100	110.70	0
T3	-	-	-	0	0	0
T4	-	-	-	0	0	0
T5	-	-	-	0	0	0
T6	-	-	-	0	0	0
	Group 7			Group 8		
	Dur sum/%	M (s)	sd (s)	Dur sum/%	M (s)	sd (s)
T1	0	0	0	72.44/94.24	72.44	0
T2	0	0	0	0	0	0
T3	44.77/23.02	44.77	0	0	0	0
T4	0	0	0	4.42/5.76	4.42	0
T5	84.82/43.65	84.82	0	0	0	0
T6	0	0	0	0	0	0
T6	64.71/33.33	64.71	0	0	0	0

Table 2.13: Durations of task positions in **task 3**. Overall duration, mean and standard deviation are expressed in seconds.

	Group 1 (%)	Group 2 (%)	Group 3 (%)	Group 4 (%)	Group 5 (%)	Group 6 (%)	Group 7 (%)	Group 8 (%)
				Task 1				
c1	9/18 (50)	26/36 (72.22)	28/30 (93.33)	4/36 (11.11)	19/33 (57.57)	7/12 (58.33)	14/20 (70)	30/38 (78.94)
c2	20/28 (71)	12/23 (52.17)	24/31 (77.41)	11/41 (26.82)	6/18 (33.330	5/8 (62.5)	3/8 (37.5)	5/6 (83.33)
c3	15/21 (71.42)	21/25 (84)	40/61 (65.57)	25/42 (59.52)	13/37 (35.13)	1/26 (3.84)	4/11 (36.36)	4/29 (13.79)
				Task 2				
c1	16/50 (32)	52/70 (74.28)	37/70 (52.85)	11/74 (14.86)	19/27 (70.37)	11/18 (61.11)	28/46 (60.86)	21/37 (56.75)
c2	28/41 (68.29)	34/49 (69.38)	41/57 (71.92)	19/93 (20.43)	3/7 (42.85)	11/12 (91.66)	13/24 (54.16)	6/16 (37.50)
c3	14/31 (45.16)	46/71 (64.78)	63/64 ()98.43	51/91 (56.04)	35/65 (53.84)	1/11 (9.09)	10/38 (26.31)	22/32 (68.75)
				Task 3				
c1	5/6 (90)	22/40 (55)	3/19 (15.78)	0/16 (0)	0	4/6 (66.66)	4/9 (44.44)	2/10 (20)
c2	16/ 20 (80)	5/10 (50)	3/11 (27.27)	2/5 (40)	0	4/4 (110)	4/14 (28.57)	0/6 (0)
c3	13/14 (92.85)	11/21 (52.38)	0/18 (0)	8/36 (22.22)	0	0/5 (0)	4/5 (80)	1/2 (50)

Table 2.14: Percentages of overlap between child positions and task positions in all groups per child and per task.

2.7 Discussion and Design Implications

The result section has shown that our measures, task performance, level of activity, type of activity, and positioning of children and task contribute to the understanding of children's collaboration on different tasks. At the same time, the measures allow us to come to conclusions about the tasks themselves. In the context of the SQUIRREL project, these conclusions allow us to inspire future scenario development. Furthermore, implications from our research can inform the design of the functional role of the robot and the specific behavior of the robot in terms of positioning in space and possible interventions to increase task engagement and enhance pro-social behaviors in playful interaction (e.g., inclusion, collaboration).

Generally, we can conclude that assembling structures from magnetic blocks turned out to be an adequate task because it enabled us to create different sub-tasks that required various abilities by the children. Also, the tasks affected the engagement as well as the social behaviors of the children. Furthermore, the task allowed us to implement sub-tasks of different levels of difficulty to suit the age and development of the children. The basic elements of the tasks (e.g., creativity, sorting, understanding shapes and objects, collaboration) can be translated to other tasks within the SQUIRREL project.

The three sub-tasks that we considered in our study were exploration, guided construction, and creative construction. The characteristics related to these tasks led to differences in children's social behavior and collaboration. Hence, when designing tasks we have to take these differences into account.

Exploration In the explorative task, the children spent most of the time in their own space working individually. Each of them examined the magnetic blocks but then also many children performed actions to attract other's attention (e.g., showing a block) to share what they had discovered. In such a task, the robot could also stay in its own space pointing out useful and interesting information about the task aiming to engage the children in the interaction. The robot should be positioned so that it is well visible for all children.

Guided Construction The guided construction task was the one that we analyzed the closest. In this task the children had a limited amount of time to perform a certain number of sub-tasks. Thus, only in this task we included performance measures to compare the groups with each other. We found big differences in performance between the groups. One explanation for these could be the age of the children. However, also groups with children of similar age, varied quite a lot in performance. Hence, in our analysis we tried to find further reasons for the differences.

One possible reason is the distribution of the roles between the children. In some groups the children were similarly active, in others one child was more active then the others and supposedly took the lead. Both strategies could work well or not. The group that performed best (group 6) had one leader and the other children supported this person (e.g., by handing over missing blocks). However, not in all groups all children accepted one as the leader. In these cases, the children did not support the other child and rather displayed aggressive behaviors, such as in group 5 that had the second-worst result in terms of performance. However, also in this group all children stayed on task - or engaged. Hence, the children seem to have the choice to reduce their efforts (or in other words to dis-engage) or to keep their efforts high and show their discomfort in other ways, such as aggressive behaviors. Hence, we should be careful to draw conclusions on the engagement level from the task performance. Also we cannot retrieve general answers to our second research question (**RQ2**: How do the non-verbal behaviors affect task engagement of all children in a group?).

The group with the worst performance was a rather balanced group in terms of activity and they performed the task in the shared space. However, they still showed a lot of aggressive behaviors and their collaboration was obviously not perfect. Hence, measures like positioning and level of activity alone do not necessarily describe the actual quality of the interaction. They have to be put in context. Hence, also with respect to our third research question (**RQ3**: How do individuals' behaviors combine to form different types of children's collaboration?) we cannot generalize.

However, the findings still provide some implications for possible robot behaviors. The robot should potentially try to stop the children from performing aggressive behaviors and to encourage exchange. In groups where the aggressive behaviors occur because one child takes the lead and the others do not accept this, the robot could try to involve all children more and to balance activity levels to reduce these tensions. It could also try to encourage collaborative actions such as exchanging blocks in the task of the study that is presented here. One possible approach to achieve this could be to move the task back into a shared space.

However, we also observed groups where children willingly clustered around the most active child by moving there or leaning over. This has two implications: the robot should position itself in a similar way and it should recognize if one child does not move close and encourage him/her to do so. Thus, the robot should promote its own and all the children's inclusion spatially.

Creative Construction The last task in our study was creative construction, where the children could assemble one shape according to their own choice. This task was designed in line with children's free play in open-ended settings rather than guided play. We included this task into our study for two reasons: Firstly, because it would give us an insight into children's non-verbal behaviors in a setting where they could take their own initiatives. This is in line with the scenario of the SQUIRREL project. Secondly, this setting would give the children the freedom to take control and determine the course of their own actions. Thus, children's behaviors are less related to the affordances of the task-related tools (e.g., images to replicate).

A close investigation of children's creative thinking was beyond the aim of this study; this is to say that we did not look at their creative task performance for this task. Rather, we focused on elements of their social interactions and the level of their activity engagement through their non-verbal behaviors. This analysis revealed that despite the fact that the children were asked to collaboratively build one construction, they tended to exhibit ego-centric behaviors and work individually. Results of the children's position and task position confirmed this tendency. Given these results, it seems that special considerations are needed for the design of tasks that support collaboration in open-ended settings and settings that support creative processes. Hence, the same might be true if a robot comes into play and we might have to work on strategies for robots to support such tasks and to engage children in more collaborative behaviors.

Limitations We want to point out some limitations of the task design, the study itself and the analysis.

One limitation with respect to the task design is that the children were sitting and assigned specific positions in the experiment area at the beginning of the session. Thus, the amount of movement in space of course was limited and we might have to test more mobile scenarios to fit the SQUIRREL project.

Another limitation of the study design was the interference of the facilitator. While we tried to avoid the impact of this as much as possible, we could not fully exclude this factor in order to ensure children's safety and understanding of the tasks.

Furthermore, the number of children that participated in this study allowed us to get some qualitative insights, however, it was not high enough to draw quantitative conclusions. Nevertheless, a deep understanding of the data in a qualitative way is a valuable step to design acceptable robot behaviors within the SQUIRREL project. Hence, in future research we aim to deepen the analysis, e.g., by including speech, i.e., what do children say and how. Due to time constraints, this analysis was out of scope for this deliverable.

Conclusion To conclude, this study was a valuable first step to understand how children collaborate on different tasks. It led to some design implications for robot behavior and inspired our research. One question that arose in this study is what roles children can take in the group. In some groups we observed that one child was rather active on the task and other children accepted this. In other groups, this did not work out at all and resulted in aggressive behaviors. We assume that one cause for this are the differences in the characters of the children. Hence, we wondered how a robot should behave in order to be accepted by the children or, in other words, what social character it should display. The study presented in the next section aims to answer this question.

3 First user study on child-robot interaction

In this section, we present the first user study on human-robot interaction that we conducted within the SQUIRREL project. It aims to give a first indication of how the robot should behave in the interaction with children and what social character it should express.

3.1 Introduction

Children appear to respond readily and strongly to robots showing social behavior [9]. Hence, child-robot interaction (CRI) is an eagerly explored area for social robots [7] in many domains (e.g. education, health care, collaborative play).

One of the main factors that contributes to initiate, sustain and maintain child-robot interactions is engagement. Engagement is a comprehensive phenomenon which comprises cognitive, affective and behavioral aspects such as focus, involvement, flow and enjoyment. Engagement in human-robot interaction has been defined in different ways, but it has been addressed mainly as '*the process by which two (or more) participants establish, maintain and end their perceived connection during interactions which they jointly undertake* [65]. Whilst some efforts to identify the features that contribute to establish and sustain the perceived connection during joint interactions have been reported in the literature, the question of which set of social behaviors a robot should be endowed with in order to initiate, maintain and support a positive engagement with a task is still to be answered.

Nevertheless, in the child-robot interaction literature one aspect that appears to be crucial in the task engagement is the style of interaction conveyed by the robot behaviors [40, 26]. The style of interaction contributes to foster children's attention, flow, and motivation on a task. At the same time it provides information about the '*character*' of the robot in the interaction [49, 22] (e.g., the robot is seen as a friend, as a teacher, as a toy).

Given the above, our first exploration of task engagement in child-robot interactions aims to shed light on the effect of *'social characters'* [1] with distinctive styles of interaction. Our study addresses the following research questions:

- **RQ1:** Do robots with different *'social characters'* (i.e., peer-like, tutor-like) expressing different styles of interaction have a different effect on the task engagement of children performing a task?
- **RQ2:** Do the *''social characters'* (i.e., peer-like, tutor-like) of the robot affect the task performance differently?

To tackle these questions, we looked into the design of robot behaviors to convey two *'social characters'*, which appear to be predominant in children task-oriented experiences, namely peers and tutors [60]. We identified sets of behaviors that reveal an *idiosyncratic* style of interaction contributing to convey a *'social character'* (i.e., peer, tutor). Following a working definition of task engagement, we evaluated the effect of the two *'social characters'*(i.e., peer-like and tutor-like) on task engagement and task performance in a Wizard of Oz (WoZ) user study.

The study entails a triadic scenario where an off-the-shelf robot (i.e., the Nao Robot) and two same sex children (6-9 years old) perform three Tangram puzzle tasks with the help of the robot.

The results contribute to gaining first insights into children's task engagement and task performance, when interacting with a social robot. Moreover, the findings set the frame for our ongoing research on designing engaging robot behaviors in a task-related contexts. Finally, the study reveals some methodological findings, which will also inform future work.

3.2 Task Engagement

Engagement is a multifaceted phenomenon, whose nature is still not well understood in human-human interaction. Given its manifold nature, investigating engagement can be challenging, as the context, the interaction and the interactant can further shape its definition [65, 57, 46].

As a result, engagement has been differently formalized across disciplines and differently approached in scientific research. Nevertheless, its multidimensional nature is widely accepted. In fact, the engagement phenomenon encompasses three dimensions or states (i.e., cognitive, behavioral, affective) [11]. The cognitive dimension/state is related to the attention processes and the inhibitory control (i.e., suppression of actions and resistance to interference from irrelevant stimuli). The behavioral dimension/state is related to the cognitive one, but it represents its motor manifestation. In other words, it comprises all the actions and fine or gross manipulations related to the cognitive processes. The affective dimension/state is related to the emotional and motivational aspect of engagement [62].

Since the early work of Sidner et al. [66], the discussions on the definition of engagement is still ongoing in human-robot interaction. Recent research highlighted that, in task related context, the social engagement (i.e., engagement with the robot) and task engagement (i.e., engagement with the task, intrinsic motivation in performing the task) should be differentiated [15, 19].

In order to investigate the effect of *'social characters'* on children's task engagement and task performance, we need to provide out own working definition of task engagement. Going a step forward from Sidner's definition, ours elaborates on insights from [14, 27, 11, 19]. It considers the level of cognitive, affective, and behavioral attributes of engagement during the interaction. These attributes (see Table 3.1) are representative of task engagement.

Attributes	Cues	Examples
Cognitive	e.g., directed gaze	paying attention to the task
Affective	e.g., smile, laughter	showing emotion in the task or report endurability, enjoyment
Behavioral	e.g., task related actions	be on task, attempting the task, perform the task as expected, completing the task

Table 3.1: The attributes of task engagement: cues and examples

The more these attributes occur in the interaction, the more motivating and involving the interaction will be. We define task engagement as:

The process by which two (or more) participants establish, initiate, maintain and end their perceived connection during task related interactions which they jointly undertake in order to accomplish a task. The initiation, maintenance and ending of the connection is regulated across the cognitive, affective and behavioral dimensions of engagement. The positive attributes of the cognitive, affective and behavioral dimensions of engagement represent the points of task engagement. The more the attributes will occur the more engaging the task will be. This definition of task engagement informed the methodology of our study.

3.3 Related Work

A wide range of aspects of child-robot engagement has been researched, mainly taking one of the following perspectives: user's engagement detection [15, 42] or identification of the elements that can lead to engagement in child-robot interactions (i.e., robot behaviors, social features, robot's styles of interaction, user modeling). Given the goal of our research, we take the latter perspective.

Most of the studies concern children with cognitive (e.g., autistic syndrome) or physical disabilities, or young users affected by chronic conditions (e.g., diabetes type I). These studies usually focus on dyadic interactions (i.e., one child interacts with the robot) [6, 25, 47, 68].

Nonetheless, some contributions have focused on normally developed children performing a task with a robot, like playing chess together, or carrying out an educational assignment [18, 36, 38, 64, 26]. The main limitation of these studies is that they account for highly specific scenarios, contexts and applications. Hence, it is challenging to trace the basic features of the child-robot engagement in task-oriented scenario. Nevertheless, one common element arises from the relevant work in the field. As a matter of fact, the literature appears to agree implicitly that the '*social character*' and the interaction style (i.e., set of communicative behaviors) of the robot plays a role in the task engagement. In a well-known field study, Kanda et al. [37]

revealed that children were more likely to engage in a learning task, when they felt that they had something in common with the robot. Hence, sharing a common ground with an equal peer robot contributed to an enhanced engagement with a task. Not only sharing commonalities, but also interacting with a specific style ascribable to character seems to play a crucial role in task engagement. For instance, Okita et al. [49] illustrated how styles of interaction influence affective engagement during a task, demonstrating that a cooperative style better supports affective task engagement and task performance. In a similar vein, Leite et al. [41] suggested that a set of emphatic behaviors, exhibited by a robot companion during a game, may encourage the child to identify the robot as a fellow peer, thus, enhancing the *endurability* with the task. A related indication emerges from the work of Belpaeme et al. [6]: a robot perceived as a peer during a game appears to be more likely to support social and task engagement.

Not only a robot *'friend'* can enhance engagement with the task, but also a robot behaving like a *'tutor'* can have a similar effect. A recent exploratory study of Kennedy et al. [40] showed how a robot tutor can enhance learning outcomes in arithmetic tasks. Despite the results on children's task engagement, the study highlighted that caution is required when applying a distinctive style of interaction to a robot *'tutor'* in a learning scenario. In fact, their findings suggest that a robot '*tutor*' endowed with sophisticated social behaviors might lead towards worse task performance, albeit sustaining the focus of attention on the learning activity.

Overall, the literature suggests implicitly that the complex set of task related and affective behaviors with a distinctive style of interaction together contribute to the definition of a robot '*character*', which appear to have an effect on task engagement. In particular, a peer *character* and a tutor one emerged as frequent robot '*characters*' in task-oriented child-robot interactions. To our knowledge, our study represents the first exploratory attempt to overtly investigate the effect of social characters with distinctive styles of interaction on children's task engagement.

Robot social characters As a working definition of *'social character'* we took the one formulated by Janlert et al. [35], who define the *'character of things'* as '*high-level attributes that help us understand and relate to them*'.

Therefore, the robot *'social character'* comprises all high-level functional, affective, communicative behaviors, which combined convey a character with social characteristics. As the robot *'social character'* is contextualized in a task-related scenario, it is further endowed with 'stylized' attributes of interaction, which together aim to lead toward task engagement (i.e., cognitive, affective and behavioral attributes) and a better task performance.

In line with the child-robot interaction and the child psychology literature, we identified two '*social characters*': peer-like and tutor like. In fact, in the interactive world of a child from six to nine years old, peers (e.g., school mates, sport mates, friends) and tutors (e.g., teachers, sport coaches) play important social roles [60].

Peer interactions are crucial at this stage of cognitive development, as they contribute to the social and emotional regulation of the child [33, 24]. Tutors represent an adult reference point outside the family playing a vital role in the cognitive development of the children, fostering attention and engagement on structured tasks [62].

The two *'social characters'* present distinctive interaction styles during task oriented activities. The interaction style of the *peer-like* character encompasses checking the understanding of the task, mutual support, emotional self-regulation, and sharing emotion [60]. The interaction style of the tutor like character is based on *scaffolding* an educational technique oriented to provide educational support to students, especially when performing tasks [53]. Based on these basic interaction styles, the two characters are further specified by behaviors in different modalities (e.g., distance, gaze, turn-taking, gestures, speech style, prosody etc.). Table 3.2 sums up the main features of the two *'social characters'*. These features guided the robot behavior design.

Interaction modalities	**Modality elements**	**Peer**	**Tutor**
vocals, paralange	pitch	high pitch	low pitch
speech	speech act, use of vocabulary	exclamations, direct speech, simple vocabulary	maieutic language, interrogative speech
proxemics	position	sitting	standing
kinesis, body language	deictic gestures	indication, sweeping	pointing and tracing
	emphatic gestures	arm in the air, suprise	head nods
	representational gestures	mocking, grasping	presenting

Table 3.2: Interaction modalities, elements and style of the robot social characters (peer and tutor)

3.4 Hypotheses

Given the definition of engagement and the related work described above, we sought to assess whether a '*social character*' with a distinctive interaction style leads children to a different level of engagement with a task. Following the literature, we defined two '*social characters*', namely a peer-like and a tutor-like character. Our specific hypotheses to address the effect of the robot *'social character'* on children's engagement are as follows:

- **H1**: There will be a significant difference in the task engagement (i.e., cognitive, affective, behavioral attributes) triggered by the two robot '*social characters*'.
- **H2**: A peer-like *'social character'* will enhance affective engagement as suggested by the literature [49]. Hence, there will be a difference in the affective dimension of task engagement.
- **H3**: A tutor-like *'social character'* will enhance the focus of attention. Hence, there will be a significant difference in the cognitive dimension of engagement triggered by the two robot 'social characters' [40].

In order to address the hypotheses, two conditions were devised:

1. *Peer-like character* ***(PC)***
2. *Tutor-like character* ***(TC)***

The two conditions are explained in the Method section.

3.5 Method

In order to investigate the effect of a robot '*social character*' on task engagement, the social character (independent variable, IV) was manipulated between-subjects. The children either interacted with a *peer-like* character (i.e., PC) or with a *tutor-like* character (i.e., TC). We were interested in how these '*social characters'* affected the task engagement. Task engagement was measured by means of behavior observations from video recordings and a questionnaire specifically attuned to children's cognitive development. The behavior observations accounted for the cognitive and behavioral attributes of task engagement, whereas the questionnaire measured the children's subjective experience related to a task (i.e., enjoyment), hence, accounting for the affective attributes of task engagement (see Table 3.1).

In the study, dyads of same gender, same age children performed three Tangram puzzle tasks with the robot exhibiting either the peer-like or the tutor-like character. In both conditions, the role of the robot was to enhance the participants' task engagement. As such, the robot was not performing the task with the children, but it was regulating the flow of the task as well as providing support and reward to the children. To this end, we designed a set of functional and affective robot behaviors for each character. These behaviors conveyed an idiosyncratic style of interaction (i.e., like a peer or like a tutor) to initiate, maintain, regulate and end the engagement on three Tangram tasks.

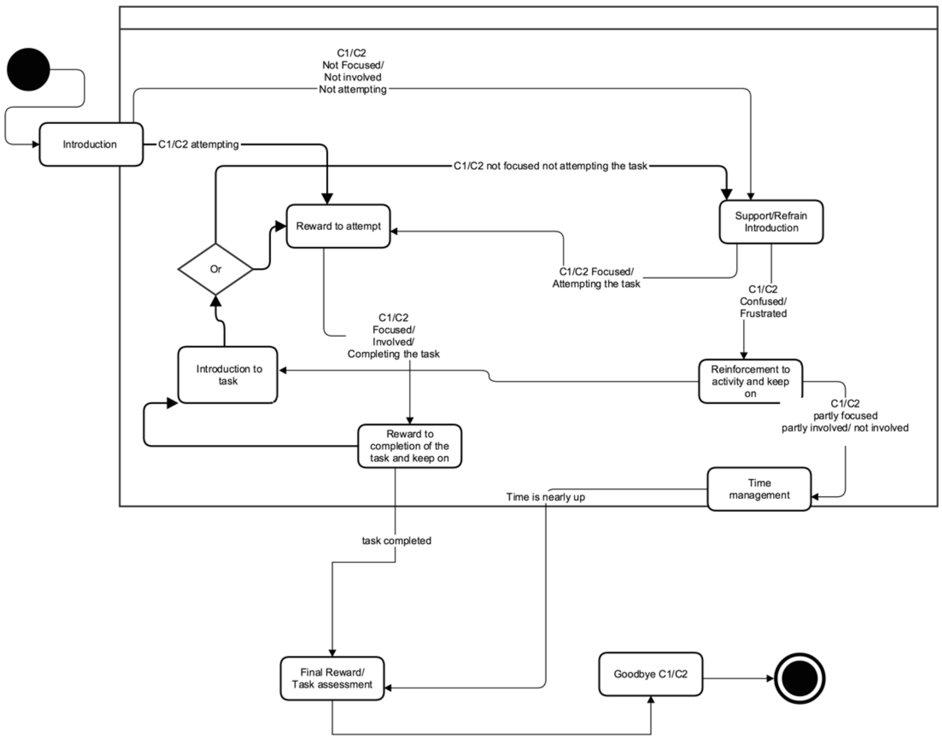


Figure 3.1: The figure presents the robot behaviors script used in the study

3.5.1 Robot and WoZ

We used the Nao robot[1] as the physical agent for the triadic scenario. The Nao is often used in child-robot interaction studies both for its features and safety [7, 83]. In the user study, the robot was remotely controlled by the researcher. A Python script controlled the communication between the remote control (i.e., input manager, the terminal) and the Nao Robot *NaoQi* [2] via Wi-Fi (i.e., output executor). The behaviors were specified with numbers 0 to 10; the list of the behaviors with the relative number were visualized in the terminal (i.e., input manager).

The researcher selected the '*social character*' (i.e., peer or tutor) and input a number to launch the behaviors in the Nao Robot. The researchers managed the transition of the behavior, i.e., the flow of the robot behaviors in the task following a script. The robot behavior script was designed to organize the behaviors in a way to regulate the task phases (e.g., introduction of the task, time management, end of the task) and the task engagement

[1]https://www.aldebaran.com/en/humanoid-robot/nao-robot
[2]http://doc.aldebaran.com/1-14/dev/naoqi/index.html

Robot Behavior	Function
Introduction Task 1	To welcome the participants and explain the first task
Refrain introduction	To repeat the information of the introduction in case of confusion
Reward to attempt	To provide reward once the task is attempted
Support	To support the participants providing suggestions and acknowledging the difficulty of the task
Reinforcement to activity and keep on	To provide positive reinforcement if the task is not accomplished and to spur to try with a new task
Reward for completion	To reward the accomplishment of the task
Introduction Task 2	To explain the second task
Introduction Task 3	To explain the third task
Final reward	To reward the participants for their participation after the third task irrespective of task completion
Time management	To tell the participants the time is almost over
Goodbye	To greet the participants

Table 3.3: Robot behaviors and their functions.

(e.g., support, reward, refrain of information). Figure 3.1 presents the robot behavior script. The behavior design is detailed in the following section.

3.5.2 Behavior design

Using Choregraphe[3], we designed and programmed twenty-two behaviors. Eleven for the PC condition and eleven for the TC condition. In order to compare the two conditions, we designed the behaviors in such a way that their function (e.g., introducing the task, providing reward, refraining information) remained the same in the two conditions.

Furthermore, each behavior in both conditions was designed employing the same interaction modalities (e.g., speech and gestures). The only difference in the design of the behaviors was the style of interaction (see Table 3.2). In other words, the way the behavior was expressed through gestures, speech, and positioning was designed either with peer-like or tutor-like characteristics (see Section 3.3.1).

The robot behaviors The behaviors were designed to (i) regulate the phases of the tasks, (ii) provide information about the state of the task

[3]https://www.aldebaran.com/en/robotics-solutions/robot-software/development

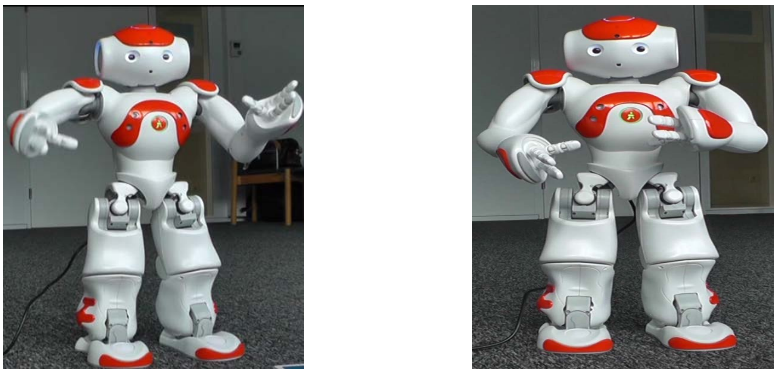
(i.e., introduction, performance, end), (iii) support the attention of the participants, (iv) provide reinforcement and support, and (v) provide reward. The resulting behaviors are described in Table 3.3 and organized in a task-related script described in Figure 3.1. Each behavior comprises the following modalities: speech, gestures, and positioning. These modalities are designed in accordance with a peer-like and a tutor-like style of interaction. In fact, exploring the literature [24, 62, 45], we identified distinctive features of speech, gestures, and positioning per condition. These features were implemented in the behaviors. The following paragraphs provide details about the distinctive features of interaction modalities per condition.

Speech: differences in TC and PC All the behaviors were supplied with speech components. We wrote speech lines consistent with the function of each behavior and with the script. The speech lines were recorded by a female researcher and added to behaviors as *.wav* files. The language of the speech lines was Dutch. In total, 116 speech lines (i.e., 58 for PC and 58 for TC) were recorded; we ensured that the total amount of speech was kept as similar as possible for both conditions.

The content of the speech lines was the same for the PC and TC conditions. In order to convey the difference in styles of interaction we modulated (i) the pitch, (ii) the syntax of the speech lines, and (iii) the semantic of the speech lines. The pitch was modulated to convey the PC voice (i.e., higher pitch) and the TC one (i.e., lower).

As for the syntax, in TC speech lines the interrogative speech in a scaffolding fashion was prevalent. In other words, in TC the robot suggests, supports and it does not instruct the child [67]. Moreover, in order to convey the tutor-pupil relationship, the robot in TC addressed the participants with the pronoun *'you'* and the noun *'kids'*. The vocabulary used the typical tutoring register which encompasses the use of simple vocabulary, but no colloquial and reinforcement expressions (e.g., '*Well, done. Keep playing like this*'). In the PC speech lines, the direct speech was prevalent and emphatic speech was frequently used. The vocabulary was simple with a colloquial twist [61]. In order to convey a sense of equal partnership, the pronoun *'we'* and the noun *'together'* were used. The robot in PC addressed the children as *'mates'*, *'buddies'* and it underlined that the activity was somehow shared [69]. Furthermore, speech lines in the PC condition exploited utterances to express affect, especially when expressing reward (e.g, *'Wow!, Nice!!!'*).

Gestures: differences in PC and TC The gestures accompanied the robot's speech. In both conditions, (i) deictic gestures, (ii) representational gestures, and (iii) emphatic gestures were used. In the TC condition, pointing and tracing gestures represented deictic and representational gestures, respectively (see Figure 3.2). These types of gestures are usually used by



(a) Tracing gesture start (b) Tracing gesture end

Figure 3.2: TC condition: Deictic and representational gestures

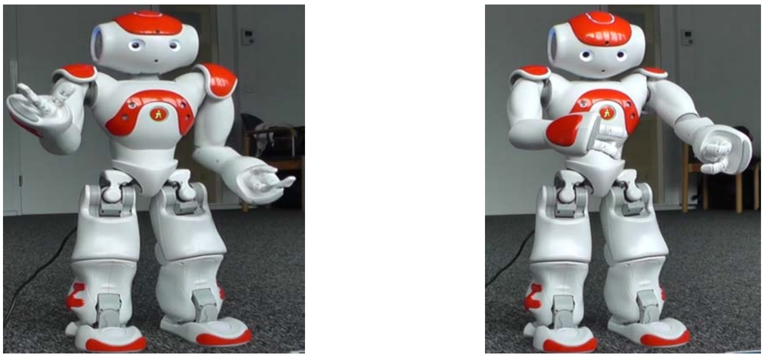
tutors and teachers to ground concepts. They are employed when explaining the task or providing help [73]. As for the emphatic gestures, head nods were used in the reinforcement and rewarding behaviors (see Figure 3.3). In the PC condition, deictic gestures like grouping were used (see Figure 3.4). Grouping is similar to presenting but it has the advantage to recall 'grasping', thus, granting the impression of the robot trying to actively engage in the task, like an equal peer. Also emphatic gestures to express perplexity, such as bringing one arm to the head, or gestures to express enjoyment like raising both arms in the air are used in the PC condition to convey the peer style of interaction [30, 3] (see Figure 3.5).

Positioning: differences in PC and TC In PC, the robot was mainly sitting with the children, again, simulating equal partnership. Moreover, the robot was close to the children [43]. In TC the robot was standing in front of the children, simulating the way a tutor would position himself/herself in space. In addition, in the TC condition the robot was positioned slightly further away from the children [62].

3.5.3 Task scenario

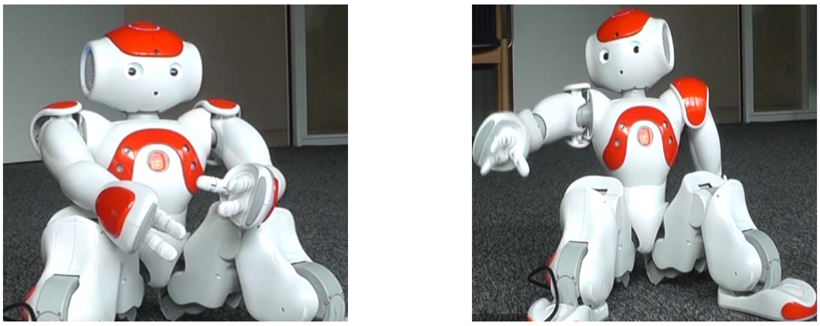
Figure 3.6 depicts the triadic scenario of the study. The child-robot interactions occurred on a playground were a couple of children (i.e., C1, C2) sitting side by side on a play-mat had to perform a puzzle task with the help of the robot, whose role was to sustain task engagement.

Following the literature on play and engagement presented in the Contextual analysis (see Section 2.2), we selected a playful sorting task for the environment. The dyad of children had to solve three Tangram puzzles of scalable difficulty. Tangram puzzles are ancient Chinese dissection puzzles consisting of seven geometric pieces, called tans, which are put together to form different shapes. The three tasks consisted in three puzzles to solve.



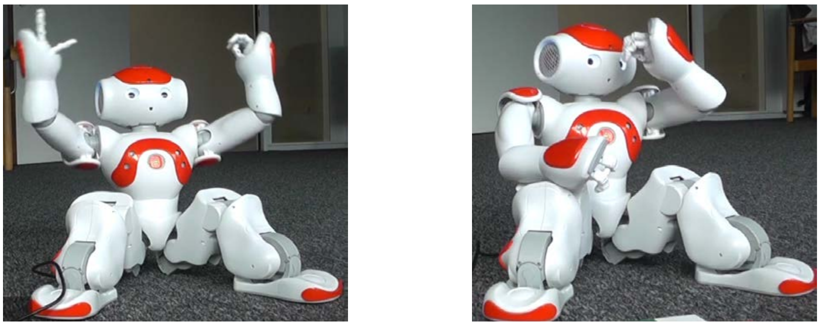
(a) Expression of reinforcement (b) Reward Gesture

Figure 3.3: TC condition: Emphatic gestures



(a) Grouping gesture (b) Mocking grasping

Figure 3.4: PC condition: Deictic and representational gestures



(a) Reward gesture (b) An expression of perplexity

Figure 3.5: PC condition: Emphatic gestures

The puzzles were presented as an outline (e.g., the perimeter of an object shape or animal shape) that had to be completed filling in the tans. The pieces to complete the puzzle were divided equally between the children, who needed to collaborate to accomplish the task. The first puzzle was designed

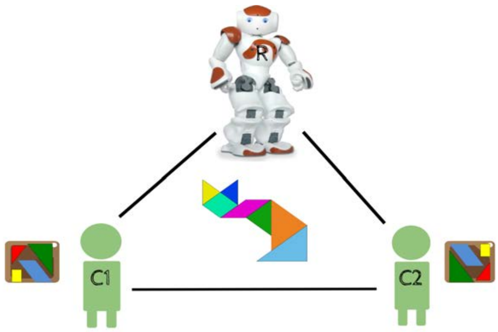


Figure 3.6: The scenario of the study

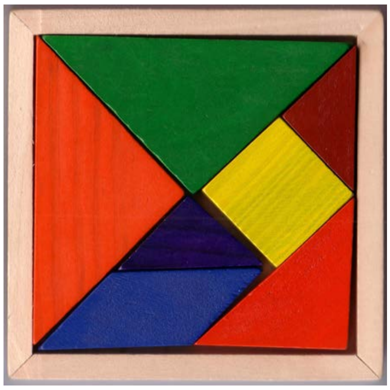


Figure 3.7: An image of one Tangram set

to be solvable for all the children, even the younger ones. It consisted of a simple outline with orientation lines, i.e., the lines defining the Tangram piece perimeter, partially completed (i.e., four puzzle pieces placed on the puzzle outline, three needed to be inserted).

The second puzzle had a medium difficulty. An outline of a house without lines was given to the children. Only a perimeter of a house was visible and one tan was placed on the outline to give a hint to the children. The children had to fill in six pieces.

The third puzzle was difficult and required creativity: the children had to organize puzzle pieces in a square without any hint. The task could be solved in a creative way, i.e., creating color patterns or mirror patterns with the pieces. In this task the children had 10 tans at their disposal; they could fill minimum 4 tans maximum 10 to complete the square puzzle.

3.6 Setup and Procedure

Setup We collected the data *in the wild* at the local school in Enschede (The Netherlands). We conducted the study in the gym room that we

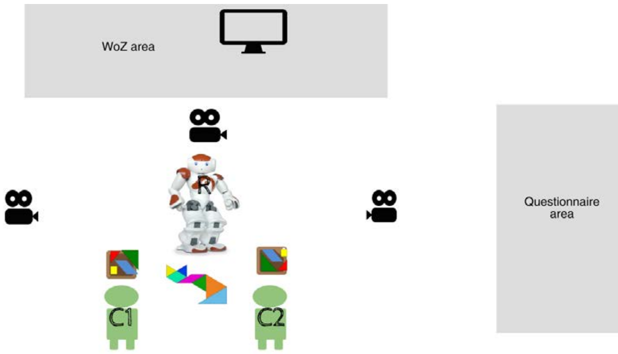


Figure 3.8: The setup of the study

divided in three areas: an experimental area, a WoZ area and a questionnaire area. The experimental area and the WoZ area were equipped with three HD cameras and computers (see Figure 3.8).

Procedure A facilitator escorted the children to the experimental area; prior to the commencement of the session, the facilitator provided an introduction to the robot in order to allow the children to familiarize with it.

The children sat on a play-mat, one next to each other; the facilitator placed the puzzle outlines and the Tangram pieces in front of them. The robot was placed opposite of the participants, outside the play-mat at a safe distance. The robot was remotely controlled from the WoZ area by one of the researchers. Although the researcher was sitting in the same room, his role was hidden from the participants. The only external actor, who had an active role in the trials, was the facilitator. The facilitator placed the puzzles' outlines for each task. The robot initiated and ended all the trials performing the '*Introduction to the first puzzle*' behaviors and the '*Goodbye*' behavior. After the interaction with the robot, the facilitator escorted the participants to the questionnaire area where the questionnaires were administered.

WoZ management The behaviors were prompted and executed following the participants' reactions. The WoZ manager was instructed to end the cycle after 21 minutes, even if the participants did not complete the tasks. Thus, for each task the WoZ manager allocated 7 minutes maximum.

Participants Twenty children, divided in ten couples (six male and four female couples), participated in the study. Their age ranged between 6 and 9 years old ($M = 7.1$, $SD = 1.28$). Following the study design, the couples

Figure 3.9: A frame from one of the recorded session. The participants are interacting with the robot during the second task.

were organized to be same sex couples (i.e., boy-boy, girl-girl). Moreover, to guarantee uniformity in the cognitive development, a further cluster concerning the age was applied, hence, the couples were the same age. Besides the age and sex requirements, the coupling was randomized: five couples were assigned to the peer character condition ($N = 10$) and the other five to the tutor character condition ($N = 10$).

3.7 Data analysis

We analyzed the videos and the questionnaires collected in the trials. We aimed to find behaviors and constructs that helped us to address our hypotheses. Following our definition of task engagement, we measured cognitive, affective, and behavioral attributes of engagement in order to investigate the effect of the robot's social characters on task engagement. To account for the cognitive attributes of engagement, we investigated focus of attention, namely the gaze behaviors of the participants. To account for the affective attributes of engagement, we designed a questionnaire on the construct of intrinsic motivation/enjoyment. To account for the behavioral attributes of engagement, we investigated task performance, rating the degree of completion of the task and analyzing task performance duration. In the following paragraphs, we provide a detailed account of the measures and the data analysis.

Behavior Observation In total we recorded approximately 126 minutes of video. All interactions were recorded with three video cameras: one capturing the central view, one the right view and one the left view. The videos were coded in the annotation tool Elan[4] following an annotation scheme

[4] https://tla.mpi.nl/tools/tla-tools/elan/

developed for the analysis which accounted for the cognitive attributes of task engagement cues (i.e., focus of attention), task performance information (i.e., efficiency and degree of completion) and contextual cues (i.e., phase of the session, phase of the task, robot behaviors). The annotations were analyzed using a Matlab toolbox called SALEM. Approximately 10% of the data (11':42") were coded by two independent coders who reached a good inter-rater reliability (Cohen's kappa: focus of attention $\kappa = 0.730$, $p = .003$, task performance $\kappa = 0.750$, $p = .000$) [31].

We annotated the focus of attention of the participants on the playground as follows: focus to the robot as gaze to the robot, focus to the task as gaze directed to the task, focus to other child, as gaze directed to either C1 or C2. We also annotated when the participants looked elsewhere outside the playground. Each cue was annotated for both children (i.e., there is a tier in ELAN for each child). These annotations were analyzed for counts and duration. As the recorded interactions differ in lengths, we normalized the results providing seconds per minute of gaze (i.e, normalization of summed duration counted gazes with overall length) and counts per minute of gaze (i.e., the rate: count normalized with overall length). These normalizations allow for comparison across interactions of different lengths. The results are compared across conditions (PC vs. TC) with two-tailed independent sample t-tests.

The task performance was evaluated in terms of completion (i.e., the degree/extent of task completion) and in terms of duration. To assess the completion, we established the following rationale: for task 1 and 2 we annotated the number of correct Tangram pieces placed on the outlines (by the participants as dyads). The scores (the number of Tangram pieces placed correctly) were respectively a maximum 3 for the first and 6 for the second task. The minimum for both tasks was 0. As for Task 3, the one which entailed some degree of creativity, we rated the results from 0 to 6 irrespective of pieces placed, as the children could complete the puzzle even without using the total number of tans available. The task was rated 6 when completed with creativity, 5 when completed, 4 when completed for three quarters 3 when half completed, 2 when completed for one quarter, 1 when attempted with few tans, 0 when no tans were inserted in the square. The degree of completion results are compared across conditions (PC vs. TC) with two-tailed independent sample t-tests.

To assess the time needed to accomplish the task, we annotated the duration of task performance as a cue for efficiency. To investigate the difference between the durations of task performances between the conditions, a Mann-Whitney U test was carried out.

Questionnaire A questionnaire was designed to measure the affective attributes of engagement (see Table 3.1). With this goal in mind, we se-

Annotation scheme		
Focus of attention	gaze to robot	*GR*
	gaze to task	GT
	gaze to other child	GC
	gaze else where	NGD
Task performance	duration of perfomance	performance
	completion	from 0 to 3
		from 0 to 6
		from 0 to 6
Phase of the session	tasks	1task
		2task
		3 task
Robot behaviors	greeting	Gr
	asking	Ask
	explaining	Exp
	reinforcing	Reinf
	supporting	Supp
	rewarding	Reward

Table 3.4: Annotation scheme used for the behavior observation. The first column contains the cognitive attributes of engagement (i.e., focus of attention), task performance information (i.e., duration and degree of completion) and contextual cues (phase of the session, robot behaviors) annotated. The central column specifies values of annotation and the last column provides the abbreviation used in Elan to perform the coding.

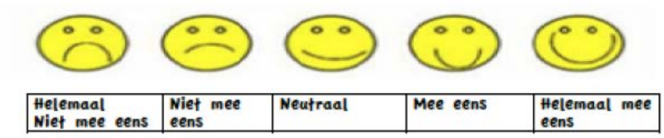


Figure 3.10: Smileyometer scale used in the questionnaire.

lected three items of the Intrinsic Motivation Inventory measuring the *Interest/enjoyment* construct [34]. The questions were translated to Dutch. Following the guidelines of the literature concerning survey methods in the field of child-computer interaction [55, 54, 12], we used the *Smileyometer*, a pictorial representation of the 5-point Likert scale anchored from *"Strongly Agree"* to *"Strongly disagree"* and we avoided reversed items. The internal reliability of the scale was 0.851 (Cronbach's alpha). Figure 3.10 presents the scale of the questionnaire. Table 3.5 presents the selected items in English.

The questionnaire presented also a closed question item. In fact, we asked the participants to say if the robot behaved more like a peer or more

IMI/Enjoyment Subscale Items
Q1 When I was playing with the tangram I was thinking on how much was enjoyable
Q2 Playing with Tangram and Nao was fun
Q3 I enjoyed the Tangram activities very much

Table 3.5: IMI/Enjoyment subscale items in the questionnaire.

like a tutor. The goal was to have an indication of the user's perception of the robot.

3.8 Results

This section presents the results for each of the conditions. In total, 125 minutes and 41,4 seconds were analyzed: 60 minutes and 47,4 seconds in PC and 64 minutes and 54 seconds in TC. The total number of annotations for the focus of attention was 1976 annotations: 1050 in PC and 926 in TC. Furthermore, we analyzed the questionnaires. The following paragraphs provide details on the findings.

3.8.1 Cognitive attributes of task engagement: focus of attention

If we look at the total of gaze behaviors directed to the playground (the gaze to the robot, to the task and to the child) the participants' gaze rate differed significantly between conditions; PC ($M = 15.479$, $SD = 0.16$) and TC ($M = 14.268$, $SD = 0.21$; $t(18) = 14.51$, $p = .0001$). However, the gaze rate attained outside the playground did not differ significantly between the two conditions; PC ($M = 1.793$, $SD = 2.265$) and TC ($M = 1.217$, $SD = 1.960$; $t(18) = 0.6081$, $p = .5507$).

In the peer condition, the participants appeared to gaze more frequently to the robot. The gaze to the robot rate was significantly higher in PC ($M = 6.218$, $SD = 0.10$) than in TC ($M = 5.28$, $SD = 1.4$; $t(18) = 2.5920$, $p = .0184$). Moreover, the participants looked at the robot significantly longer (i.e., gaze seconds/per minutes) in PC ($M = 6.190$, $SD = 1.6$) than TC ($M = 4.480$, $SD = 0.41$; $t(10) = 3.2739$, $p = .0084$).

As for the gaze focused on the task, the rate of gaze was significantly higher in PC ($M = 6.43$, $SD = 0.16$) than in TC ($M = 5.023$, $SD = 0.21$; $t(18) = 16.8530$, $p = .0001$). The average amount of gaze seconds per minutes on the task was also significantly different in the two conditions (PC: $M = 3.09$, $SD = 0.040$; TC: $M = 2.04$, $SD = 0.08$; $t(18) = 8.1398$, $p = .001$).

We also looked at the focus of attention (i.e., gaze) directed to the other child. No significant difference in the rate of gaze to the other child was found between PC ($M = 2.829$, $SD = 0.63$) and TC ($M = 2.742$, $SD = 0.38$;

Task performance: Efficiency	*min*	*max*	*M*	*SD*
PC condition				
Perf T1	24.10	63.50	38.12	17.23
Perf T2	30.47	165.10	87.97	63.61
Perf T3	40.40	276.90	124.23	99.68
Tot. Perf	24.10	276.90	83.43	73.58
TC Condition				
Perf T1	49.50	179.80	95.17	58.93
Perf T2	49.00	170.00	126.76	53.64
Perf T3	67.00	301.80	211.84	101.06
Tot. Perf	49.00	301.80	132.16	84.07

Table 3.6: Task performance: efficiency expressed in performance duration *(s)*. The table presents the min/max duration, mean and standard deviation of total duration of performance (Tot. Perf), duration perfomance in the first task (Perf T1) second task (Perf T2), third task (Perf T3).

$t(18) = 0.3739$, $p = .1184$). In a similar vein, the average amount of gaze seconds per minute were also not statistically significant across conditions (PC: $M = 5.170$, $SD = 1.08$, TC: $M = 6.04$, $SD = 8.43$; $t(18) = 0.3237$, $p = 0.7499$).

3.8.2 Affective attributes of engagement: enjoyment

Overall, both in the peer character condition and the tutor character condition the results were polarized towards the highest score. The children reported high enjoyment with the robot and the task in both conditions. A Shapiro-Wilk test on the data confirmed the non-normality of the distributions with an evident negative skew. Given the distribution of the data and the presence of outliers, a Mann-Whitney U test was carried out to determine if there were differences in the IMI/Enjoyment score between PC and TC. Enjoyment scores for PC ($meanrank = 10.30$) and TC ($meanrank = 10.70$) were not statistically significantly different ($U = 48$, $Z = -154$, $p = .878$).

3.8.3 Behavioral attributes of task engagement: task completion

Task 1 was completed by all the participants in both conditions. Task 2 was completed by 80% of the participants (4 out of 5 dyads) in the peer character condition and just by 20% of participants (1 out of 5 dyads) in the tutor character condition. Likewise, Task 3 was completed by 80% of the participants in the peer character condition and by only 20% of participants in the tutor character condition. An independent-samples t-test was carried out to determine if there were differences in participants' effectiveness between the group who performed Task 2 in PC or TC. The task performance was better in PC ($M = 5.80$, $SD = 0.44$) than in the TC ($M = 3.80$, $SD = 2.16$), but no statistically significant difference was found (95% CI, $t(8) = 2.020$, $p = .078$). As for Task 3, the task performance was better in PC ($M = 5.60$, $SD = 0.89$) than in TC ($M = 2.80$, $SD = 0.83$) and a statistically significant difference was found ($t(8) = 5.112$; $p = .001$).

Therefore, H1 was only supported for the third task, the most difficult one, but not for the first and second task.

3.8.4 Behavioral attribute of task engagement: efficiency

We also annotated the duration of task performance as a cue for efficiency (i.e., time needed to perform the task). The results also revealed that the participants in TC not only had difficulties to accomplish the tasks, but they also took more time to perform the tasks than the participants in PC. Table 3.6 depicts the descriptive statistics relative to the total duration of task performances and the duration of task performance for the first, second, third task.

A Mann-Whitney U test conducted on the total performance duration confirmed that the participants in the tutor condition ($meanrank = 18.50$) took significantly more time ($U = 154$, $Z = 2.139$, $p = .033$) to perform the task than the participants in the peer condition ($meanrank = 11.73$), thus supporting H1.

3.8.5 Children's perception of the robot: peer or tutor

The majority of the children perceived the robot as a friend, hence, even when they were assigned to the tutor condition they saw the robot as a peer; only 15% of the children perceived the robot as a tutor/teacher, 85% of the participants experienced the robot like a fellow peer. In TC, 80% of the participants identified the robot as a peer, while the remaining 20% of the participants correctly identified the '*social character*' of the robot (i.e., tutor). In PC, the results appeared reversed as 90% of the participants correctly identified the peer-like *social character* of the robot, while only 10% perceived it as a tutor.

3.9 Discussion

In our first user study we presented an exploration of the effect of two *'social characters'*, peer-like and robot-like, on children's task engagement. From the analysis of the results of cognitive attributes of engagement (i.e., focus of attention), affective attributes of engagement (i.e., intrinsic motivation, enjoyment) and behavioral attributes of engagement (i.e., task performance) we can account for an effect of the two '*social characters'* on cognitive-behavioral attributes of task engagement, but not for an effect on affective engagement, thus H1 is partially supported. Despite the results partially supporting H1, the children did not seem to perceive a difference between the '*social characters*', as they reported the robot being a peer in both conditions.

If we look at each attribute independently, the peer-like '*social character*' appears to have an effect on the focus of attention of the children, triggering significantly more focus towards the robot and the task than the tutor-like '*character*'. In other words, the peer-like '*social character*' attracted more attention towards itself, its explanations, reinforcements, rewards and towards the task. These results contradict H3 and the literature, suggesting that a social tutor enhances the attention of children on a task and on the tutor's explanations [40]. One possible reason for this unexpected outcome is that the style of interaction exhibited in the peer-like condition, friendly, equal, direct, emphatic is more suitable for a playful task like the Tangram tasks, than for a more scaffolding-oriented one.

This explanation can be applied also on the results of the behavioral attributes of task engagement. In fact, our findings showed higher task performance and task effectiveness in the peer-like '*social character*' condition in the more difficult tasks. This supports the outcome on the focus of attention, thus, indicating that the style of interaction of a peer-like '*social character*' as promoting an enhanced cognitive-behavioral task engagement.

On the other hand, the results on affective attributes of task engagement discard the possibility of a significant difference of the effect of '*social characters*' on children's enjoyment. In contrast with the literature [41], the peer-like '*social character*' is not providing an enhanced effect on affective attributes of engagement, thus H2 is not supported. Despite performing differently and in some cases worse, the children scored their enjoyment very high in both conditions, thereby not accounting for less enjoyment in case of poor task performance. One explanation of these results could be a '*novelty effect*' of the robot on the children. In fact, as this user study represents the very first encounter with a robot for the children, their perception of the robot and of the task might have been biased by their curiosity and excitement for the new experience. This might have led to less awareness on the perceived enjoyment.

Furthermore, the results on affective attributes of engagement and the open question results could be suspect to the *suggestibility effect* [79]. In fact, it has been reported by the literature that children's self-reporting can be affected by the desire to please the facilitator and by the pressure of the school context, leading to score the questionnaire in a positive way to not disappoint the researcher. The results account also for a polarization effect, the children tend to polarize their responses to the questionnaire to the extreme anchors [12]. This effect might indicate that their capability to discern a scale is not refined yet, suggesting the inappropriateness of survey methods with this user group.

Another aspect reported by the literature, that might have contributed to the questionnaire's results, is the expectation of the children towards the robot [7]. In fact, the appearance of the Nao robot, with his friendly and pleasant embodiment could have increased the expectations of bonding, thus contributing in seeing this robot like a fellow enjoyable peer.

Overall, our first investigation of children's engagement highlights once more the complexity of the engagement phenomenon and the necessity to look into its cognitive, affective and behavioral attributes, as they contribute differently to task engagement. With respect to our research questions (RQ1 and RQ2) on the effect of peer-like and tutor-like '*social characters*' on task engagement, our user study indicates an effect of the peer-like character with its distinctive interaction style on cognitive-behavioral task engagement (i.e., focus of attention and task performance). Moreover, from this first exploration we draw some methodological conclusions, namely the inappropriateness of survey methods with this user group and the need to address children's expectation towards a social robot.

3.10 Limitations and future work

Our first user study suffered some limitations, which will be addressed in future work.

Firstly, the sample size attenuates the results, as the study was carried out with only 20 participants.

Secondly, we are aware that our findings on behavioral attributes of engagement cannot provide a comprehensive account of the task performance results, as they do not account for children's prior level of ability on the task. In our experimental design, we tried to overcome this limitation by addressing the possible discrepancies in the children's cognitive development. This is why we favored uniform age and same gender dyads. Nevertheless, we are aware that the task performance results need to be grounded by each child's sorting ability level and we will add this aspect to our future study designs. In the long run, this aspect will be added *en route* to endow the robot with the capability to adapt to the users' abilities. Similarly, we are aware that, to provide a complete account of the task engagement dynamics, future work should take into account the effect of task engagement on the small group of children, hence discriminating the results per child.

Furthermore, we are aware that our results should be compared with a human peer and human tutor interactions to be further grounded. This was not possible for this very first exploration, but it will possibly be addressed in follow-up studies.

Lastly, our investigation looked into the design of engaging child-robot interaction from a *top-down* perspective. We looked at the overall effect of a '*social character*' with a distinctive style of interaction. Our goal was to have early indications on what '*social character*' a robot should display when it inserts itself in a play-group. As a result, we did not look into the effect of single robot behaviors and single style interaction features. This was beyond the scope of the study. In light of our findings on the effect of a robot's '*social character*' on task engagement, we will proceed with *bottom-up* investigations to discern low-level engaging behaviors and interaction style features oriented at promoting task engagement.

4 Corpus of data from first user study

In the context of the user study (see section 3), we recorded our first child-robot interaction corpus. The corpus comprises ten videos and ten richly annotated files; each video accounts for one experimental session and each annotated file represents the output of the annotations realized with the annotation tool ELAN.

The following sections detail the terms of use of the corpus, the recorded videos and the annotated files.

4.1 Use of the corpus: ethical consideration

In this section, we provide details about the terms of use of the corpus. These terms of use affect: (i) the consultation, (ii) further data analysis and (iii) the dissemination of the corpus by project partners and third parties.

As explained in Section 3, participants were recruited at a local elementary school, partner of the University of Twente. The approval to participate in the study was given by the parents of the children by written agreement on the basis of an information brochure and a consent form.

The consent form specified the usage of the collected data in terms of anonymity, personally identifiable information and information dissemination. The parents could opt for giving a full consent, thus granting permission to participate in the user study, to allow audiovisual recordings and to disseminate the data for research purposes. Otherwise the parents could opt for giving a partial consent, not granting permission to disseminate the audiovisual data.

In total, fifteen children had full consent and five children had partial consent (see Table 4.1 for details). As a consequence, researchers who would like to access the corpus should bear in mind the following terms of use:

- Data obtained from the user study is not disclosed to third parties. The corpus is available to the SQUIRREL project partners for research purposes only.
- Video screenshots and video recordings must be anonymized if disclosed in authorized academic dissemination (e.g. academic publications).
- Participants with partial consent cannot be included in any dissemination material, but their data can be analyzed by SQUIRREL project partners.
- The researcher must operate in accordance with privacy legislation, hence, data in which subjects are identifiable is carefully stored and is deleted whenever the interest of the research allows for this.

File	Consent
C1C2S1PF.mov/.eaf	Full
C1C2S2PM.mov/.eaf	Full
C1C2S3PM.mov/.eaf	Full
C1C2S4PF.mov/.eaf	Partial
C1C2S5TM.mov/.eaf	Partial(1 participant)
C1C2S6TF.mov/.eaf	Partial(1 participant)
C1C2S7TM.mov/.eaf	Full
C1C2S8TM.mov/.eaf	Full
C1C2S9TF.mov/.eaf	Full
C1C2S10PM.mov/.eaf	Partial(1 participant)

Table 4.1: The table lists the files of the corpus, both videos and ELAN files. In addition, it provides information about the consent.

In other words, the corpus can be consulted and used by SQUIRREL project partners for research purposes (e.g., data analysis, internal presentations), but it cannot be available to the academic community or the general public.

Moreover, in case of authorized dissemination (e.g, academic publications, presentations at conferences, symposiums) video screenshots or video recordings should not be identifiable, hence the visual features of the participants must be unrecognizable.

Overall, researchers are obliged to the above mentioned terms of use in order to conform to the ethical principles of this research.

4.2 Videos

In this section we provide details about the videos collected during the first user study. As explained in Section 3 the interaction between the participants and the robot were audio and video recorded.

Format and video synchronization. The audiovisual recordings captured three views: (i) a central view were the participants can be seen frontally, (ii) a right view, which capture the interactions from the right side and (iii) the left view, which capture the interactions from the left side.

In total, thirty raw videos were captured, one from each camera angle. The raw videos were processed using Final Cut Pro3 and converted in a standard format (i.e., *.mov*). Subsequently, the videos were synchronized as follows: all the three camera angles start with the children sitting on the mat opposite to the robot waiting for the first behavior of the script to start, namely *'Introduction to the first puzzle'* and end with the last behavior of the script, namely *'Goodbye'*. Thus, the above mentioned anchor points

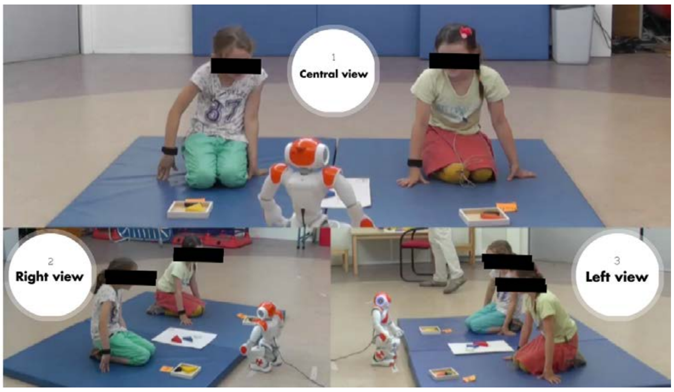


Figure 4.1: The figure depicts a frame of the videos after synchronization and split screen processing. The central, right and left views are labeled accordingly.

were taken as referents for the time aligned synchronization. The former was taken at starting point, the latter as ending point.

In addition, the three video streams are merged together in one video. In fact, using the multi-camera function of the video processing tool, the videos were merged in a single stream. The resulting ten videos present the three camera angles in a split-screen-like manner. In other words, the videos present all the streams synchronized in one (see Figure 4.1).

Use of the videos. The ten videos recorded during the first user study can be used for further analysis in accordance with the general terms of use discussed in the above section. The video format is compatible with the most common annotation software (e.g., ELAN, ANVIL[5]) and data analysis tools (e.g., MATLAB). Upon request, the master files of the frontal view videos can be shared with the SQUIRREL project partners.

4.3 ELAN Files

The corpus includes ten richly annotated files. These files represent the output of the behavior observations of the first user study.

Format and description of the files. As explained in Section 3, the video data were annotated in ELAN following an annotation scheme (see Table 3.4 used to analyzed participant's task engagement. Details of the annotation scheme can be found in Section 3. ELAN (EUDICO Linguistic Annotator)[6] is a free annotation tool that allows to create time aligned annotation for multimodal data.

[5]http://www.anvil-software.org/
[6]https://tla.mpi.nl/tools/tla-tools/elan/

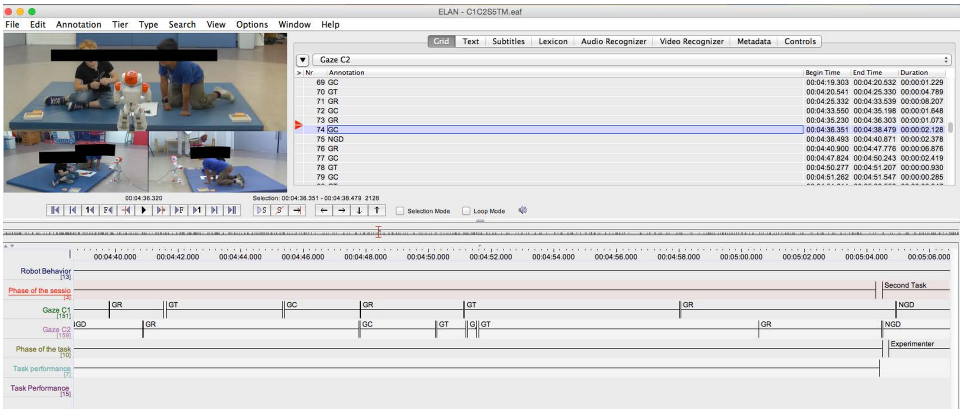


Figure 4.2: The figure presents a screenshot taken from one of the ELAN files.

The tool revolves around the concept of tier, namely a set of annotations that share the same characteristics. Each tier has a controlled vocabulary, namely a list of possible annotation values. In the ELAN files of the corpus, the tiers and the annotation values are based on the user study's annotation scheme (see Table 3.4). As a result, the ELAN files included in the corpus contain all the information about tiers, annotations values, time alignments and links to media files, e.g., the videos (see Figure 4.2).

The format of the ELAN files is *.eaf*, a tool-specific extension. Hence, the files can be read only in ELAN.

Use of ELAN files. Once read in ELAN, the files can be further analyzed, adding more tiers and annotation values. Moreover, the tool allows the integration of third part tools (e.g., Praat).

In order to analyze the annotations the files can be analyzed manually in the tool or, alternatively, it is possible to export the files in text or XML files. Once exported, in order to perform statistical analysis, these files have to be imported in data analysis software (e.g., SPSS, Matlab). An alternative to this laborious procedure is SALEM (Statistical AnaLysis of Elan files in Matlab)[31], a Matlab tool developed by the University of Bielefeld. The tool is free and available for research purposes. As already explained in Section 3, SALEM was used to analyze the corpus.

References

[1] Alac, M., Movellan, J., and Tanaka, F. When a robot is social: Spatial arrangements and multimodal semiotic engagement in the practice of social robotics. *Social Studies of Science 41*, 6 (Oct. 2011), 893–926.

[2] Alderson, P., and Morrow, V. *The Ethics of Research with Children and Young People.* Sage, London, 2011.

[3] Alibali, M. W., Flevares, L. M., and Goldin-Meadow, S. Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology 89*, 1 (1997), 183–193.

[4] Amabile, T. M. *The Social Psychology of Creativity.* Springer-Verlag, New York, 1983.

[5] Bandura, A. *Social Foundations of Thought and Action: a Social Cognitive Theory.* Prentice Hall, Upper Saddle River, NJ, 1986.

[6] Belpaeme, T., and Al. Multimodal Child-Robot Interaction: Building Social Bonds. *J. of Human-Robot Interact.* (Jan. 2013), 33–53.

[7] Belpaeme, T., Baxter, P., Greeff, J. D., Kennedy, J., Read, R., Looije, R., Neerincx, M., Baroni, I., and Zelati, M. C. Child-Robot Interaction : Perspectives and Challenges. *Social Robotics* (2013), 452–459.

[8] BERA. Ethical Guidelines for Educational Research. Tech. rep., London, UK, 2011.

[9] Bethel, C. L., Stevenson, M. R., and Scassellati, B. Secret-sharing: Interactions between a child, robot, and adult. *2011 IEEE International Conference on Systems, Man, and Cybernetics* (Oct. 2011), 2489–2494.

[10] Blakemore, S.-J. The developing social brain: imprications for education. *Neuron 65* (2010), 744–747.

[11] Blumberg, C., Eds, S. M. F., Context, D. G. A., Deater-deckard, K., Chang, M., and Evans, M. E. Engagement States and Learning from Educational Games. 21–30.

[12] Borgers, N. Children as Respondents in Survey Research: Cognitive Development and Response Quality.

[13] BPS. Code of Human Research Ethics. Tech. rep., Leicester, 2011.

[14] Brien, H. L. O., and Toms, E. G. What is User Engagement ? A Conceptual Framework for Defining User Engagement with Technology. *Journal of the American Society for Information Science and Technology 59*, 6 (2008), 938–955.

[15] Castellano, G., Leite, I., Pereira, A., Martinho, C., Paiva, A., and McOwan, P. W. Detecting Engagement in HRI: An Exploration of Social and Task-Based Context. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing* (Sept. 2012), 421–428.

[16] Christensen, P. H. Children's participation in ethnographic research: Issues oh power and representation. *Children and Society 18*, 2 (2004), 165–176.

[17] Cohen, L., Manion, L., and Morrison, K. *Research Methods in Education*, vol. 55. Routledge, 2007.

[18] Cooney, M. D., Kanda, T., Alissandrakis, A., and Ishiguro, H. Interaction design for an enjoyable play interaction with a small humanoid robot. *2011 11th IEEE-RAS International Conference on Humanoid Robots* (Oct. 2011), 112–119.

[19] Corrigan, L. J., Peters, C., and Castellano, G. Social-Task Engagement: Striking a Balance between the Robot and the Task. In *Embodied Commun. Goals Intentions Work ICSR'13* (2013), pp. 1–7.

[20] Corsaro, W. *The sociology of childhood.* Pine Forge Press, Thousand Oaks, California, 1997.

[21] Craft, A., Cremin, T., Burnard, P., and Chappell, K. Developing creative learning through possibility thinking with children aged 3-7. In *Creative learning 3-11 and how we document it.* Trentham, London UK, 2007.

[22] Díaz, M., Nuño, N., Sàez-pons, J., Pardo, D., Angulo, C., and Andrés, A. Building up child-robot relationship from initial attraction towards social engagement. *Automatic Face & Gesture Recognition and Workshops.IEEE International Conference on. IEEE 27* (2010), 1–6.

[23] E, S. A. Communication and regulation in psychotherapy. *Psychiatry 26* (1963), 126–136.

[24] Fawcett, L. M., and Garton, A. F. The effect of peer collaboration on children's problem-solving ability. *The British journal of educational psychology 75*, Pt 2 (June 2005), 157–69.

[25] Feil-seifer, D. Robot assisted theraphy for children with Autism Spectrum Disorders. *Proceedings of the 7th international conference on Interaction design and children. ACM* (2008), 49–52.

[26] Fink, J., Dillenbourg, P., Wille, F., and Franinovi, K. Which Robot Behavior Can Motivate Children to Tidy up Their Toys ? Design and Evaluation of Ranger. *9th ACM/IEEE International Conference on Human-Robot Interaction (HRI" 14)* (2014).

[27] Fredricks, J. A., and Blumenfeld, P. C. School Engagement : Potential of the Concept , State of the Evidence. *Rev. Educ. Res.* (2009), 59–109.

[28] Froebel, F. D. Appleton, New York, 1909.

[29] Gibson, W. The theory of affordances. In *Perceiving, Acting, and Knowing: Toward and Ecologival Psychology*, R. Shaw and J. Bransford, Eds. Erlbaum, Hillsdale, NJ, 1977, pp. 67–82.

[30] Goldin-Meadow, S., and Sandhofer, C. M. Gestures convey substantive information about a child's thoughts to ordinary listeners. *Developmental Science 2* (1999), 67–74.

[31] Hanheide, M., Lohse, M., and Dierker, A. SALEM - Statistical AnaLysis of Elan files in Matlab. In *Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality* (2010), pp. 121–123.

[32] Holzman, L. Without Creating ZPDs There is No Creativity. In *Vygotsky and Creativity: a Cultural-historical Approach to Play, Meaning Making and the Arts*, C. Connery, V. John- Steiner, and M. Majanovic-Shane, Eds. Peter Lang Publishers, New York, 2010.

[33] Hubbard, J. A., and Coie, J. D. *Emotional Correlates of Social Competence In Children's Peer Relationships.* 1994.

[34] Intrinsic Motivation Inventory. Intrinsic Motivation Inventory (IMI). Tech. rep., 1994.

[35] Janlert, L.-E., and Stolterman, E. The character of things. *Design Studies 18* (1997), 297–314.

[36] Janssen, J. B., Wal, C. C. V. D., Neerincx, M. A., and Looije, R. Motivating Children to Learn Arithmetic with an Adaptive Robot Game. In *Social Robotics.* Springer Berlin Heidelberg, 2011, pp. 153–162.

[37] Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human-Computer Interaction 19*, 1 (June 2004), 61–84.

[38] Kanda, T., Sato, R., Saiwaki, N., and Ishiguro, H. A two-month field trial in an elementary school for long-term humanrobot interaction. *Robotics IEEE Trans* (2007), 962–971.

[39] Kangas, M. Creative and playful learning: Learning through game co-creation and games in a playful learning environment. *Thinking Skills and Creativity 5*, 1 (2010), 1–15.

[40] Kennedy, J., Baxter, P., and Belpaeme, T. The Robot Who Tried Too Hard : Social Behaviour of a Robot Tutor Can Negatively Affect Child Learning. In *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Intereaction* (2015), p. to appear.

[41] Leite, I., Castellano, G., Pereira, A., Martinho, C., and Paiva, A. Empathic Robots for Long-term Interaction: Evaluating Social Presence, Engagement and Perceived Support in Children. *International Journal of Social Robotics 6* (2014), 329–341.

[42] Leite, I., Mccoy, M., Ullman, D., Salomons, N., Scassellati, B., and Haven, N. Comparing Models of Disengagement in Individual and Group Interactions. In *10th ACM/IEEE International Conference on Human Robot Interaction Conference (to appear)* (2015).

[43] Lomranz, J. F. Children's Personal Space As a Function of Age and Sex. *Developmental psychology 11*, 5 (1975), 541–545.

[44] Marshall, C., and Rossman, G. B. *Designing Qualitative Research.* Sage, Newbury Park, CA, 1995.

[45] Merola, G., Poggi, I., Psicologia, D., and La, R. Multimodality and Gestures in the Teachers Communication Multimodality in the Classroom A Human s Meanings. *Gesture-based communication in human-computer interaction* (2004), 101–111.

[46] Michalowski, M., Sabanovic, S., and Simmons, R. A spatial model of engagement for a social robot. In *9th IEEE International Workshop on Advanced Motion Control, 2006.* (2006), Ieee, pp. 762–767.

[47] Michaud, F. Y., Salter, T. Y., Duquette, A. ., Mercier, H. ., Lauria, M. Y., Larouche, H. ., and Larose, F. . Assistive Technologies and Child-Robot Interaction. *American Association for Artificial Intelli gence ii* (2007), 8–9.

[48] O' Brien, H. L., and Toms, E. G. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Scuence and Technology 59*, 6 (2008), 938–955.

[49] Okita, S., and Al. Multimodal approach to affective human-robot interaction design with children. In *ACM Trans. on Interact. Intell. Syst.* (Oct. 2011), vol. 1, pp. 1–29.

[50] Papert, S. Papert on Piaget. *Time: The Century's Greatest Minds* (1999), 105.

[51] Pellegrini, A. D. *The Role of Play in Human Development.* Oxford University Press, New York, 2009.

[52] Piaget, J. *The Origins of Intelligence in Children.* International University Press, New York, 1952.

[53] Poggi, I., Errico, F. D., and Leone, G. Pedagogical Stance and its multimodal expression. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems* (2003), no. 1976, pp. 3233–3240.

[54] Read, J., Macfarlane, S., and Casey, C. Endurability , Engagement and Expectations : Measuring Children s Fun. In *Interaction Design and Children* (2002), vol. 2, pp. 1–23.

[55] Read, J. C. Understanding Childrens Interactions : Evaluating Childrens Interactive Products. In *Emerging Approaches to Research and Design Practice-Interactions* (1999), pp. 26–29.

[56] Resnick, M. All I Really Need to Know (About Creative Thinking) I Learned (By Studying How Childre Learn) in Kindergarten, 2007.

[57] Rich, C., Ponsler, B., Holroyd, A., and Sidner, C. L. Recognizing engagement in human-robot interaction. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Mar. 2010), Ieee, pp. 375–382.

[58] Robson, S., and Rowe, V. Observing young children's creative thinking: engagement, involvement and persistence. *International Journal of Early Years Education 20*, 4 (2012), 349–364.

[59] Rogoff, B., and Damon, W. Cognition as a Collaborative Process. In *Handbook of Child Psychology*, D. Kuhn and R. S. Siegler, Eds., vol. 2. Wiley, New York, 1998, pp. 679–744.

[60] Rubin, K. H., Bukowksi, W. M., and Laursen, B. *Handbook of Peer Interactions, Relationships, and Groups.* Guilford Press, New York, New York, USA, Aug. 2011.

[61] Sachs, J., and Devin, J. Young children's use of age-appropriate speech styles in social interaction and role-playing, 1976.

[62] Sandra, C., Reschly, A. L., and Wylie, C. *Handbook of research in Student Engagement.* Springer, New York, New York, USA, 2012.

[63] Scardamalia, M., and Bereiter, C. Knowledge building: Theory, pedagogy, and technology. In *Cambridge Handbook of the Learning Sciences*, K. Sawyer, Ed. Cambridge University Press, New York, 2006, pp. 97–118.

[64] Shahid, S., Krahmer, E., and Swerts, M. Childrobot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend? *Computers in Human Behavior 40* (2014), 86–100.

[65] Sidner, C. L., Kidd, C. D., Lee, C., and Lesh, N. Where to Look : A Study of Human-Robot Engagement. In *Proceedings of the 9th international conference on Intelligent user interfaces* (2004), ACM, pp. 74–84.

[66] Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N., and Rich, C. Explorations in engagement for humans and robots. *Artificial Intelligence 166*, 1-2 (Aug. 2005), 140–164.

[67] Singer, M. a., and Goldin-meadow, S. Children Learn When Their Teacher s Gestures and Speech Differ. *Society 16*, 2 (2005), 85–89.

[68] Tapus, A., Peca, A., Aly, A., Pop, C., Jisa, L., Pintea, S., Rusu, A. S., and Daniel O. David. Children with autism social engagement in interaction with Nao, an imitative robotA series of single case experiments. *Interaction studies 13 3* (2012), 315–347.

[69] Teasley, S. D. The role of talk in children's peer collaborations. *Developmental Psychology 31*, 2 (1995), 207–220.

[70] Tomasello, M. *A natural history of human thinking.* Harvard University Press, Cambridge, MA, 2014.

[71] Torrance, E. P. *Education and the creative potential.* University of Minnesota Press, Minneapolis, MN, 1963.

[72] Uttal, D. H., and Cohen, C. A. Spatial Thinking and STEM Education: When, Why, and How. *Psychology of Learning and Motivation 57* (2012), 147–181.

[73] Valenzeno, L., Alibali, M. W., and Klatzky, R. Teachers' gestures facilitate students' learning: A lesson in symmetry, 2003.

[74] Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., and Chang, A. Deconstructing

Building Blocks: Preschoolers' Spatial Assembly Performance Relates to Early Mathematical Skills. *Child Development 85*, 3 (2014), 1062–1076.

[75] Vygotsky, L. S. The role of play in development. In *Mind in society: The development of higher psychological processes*, M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, Eds. Harvard University Press, Cambridge, MA, 1978, pp. 92–104.

[76] Wai, J., Lubinski, D., and Benbow, C. P. Spatial Ability for STEM Domains: Aligning Over 50 Years of Cumulative Psychological Knoweledge Solidifies its Importance. *Journal of Educational Psychology 101*, 4 (2009), 817–835.

[77] Wallas, G. The Art of Thought. In *Creativity*, P. E. Vernon, Ed. Penguin, Harmondsworth, UK, 1970.

[78] Wang, M.-T., and Degol, J. Staying Engaged: Knowledge and Research Needs in Student Engagement. *Child Development Perspectives 8*, 3 (2014), 137–143.

[79] Warren, A. R., and Marsil, D. F. Why children's suggestibility remains a serious concern. *Law and Contemporary Problems 65*, 1993 (2002), 127–147.

[80] Wiggins, G. A., and Bhattacharya, J. Mind the gap: an attempt to bridge computational and neuroscientific approaches to study creativity. *Frontiers in human neuroscience 8* (2014), 540–577.

[81] Wing-yi Cheng, R., Lam, S.-f., and Chung-yan Chan, J. When high achievers and low achievers work in the same group: The roles of group heterogeneity and processes in project-based learning. *British Journal of Educational psychology 78*, 2 (2008), 205–221.

[82] Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., and Sloetjes, H. ELAN: A professional framework for multimodality research. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC)* (2006), pp. 1556–1559.

[83] Xu, J., Broekens, J., Hindriks, K., and Neerincx, M. a. Effects of bodily mood expression of a robotic teacher on students. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Sept. 2014), Ieee, pp. 2614–2620.