

DELFT UNIVERSITY OF TECHNOLOGY

MASTER THESIS  
MSc CS THESIS PROJECT EEMCS

---

**Exploring Children's Choices in an  
Educational Game on Neurodiversity:  
Reveiling Underlying Values through  
Robot's Socratic Questioning**

---

*Author:*

Laura Ottevanger, 4584953

*Supervisors:*

Catharine Oertel, Mark Neerincx

September 12, 2023



## **Abstract**

In the Netherlands, there is a shortage of primary school teachers, due to this shortage, teachers often do not have a lot of one-on-one time with the students. A social robot could be the solution to creating more one-on-“one” time with the students. In addition to relieving some workload for the teachers, a social robot could assist in gaining insight into personal values. Gaining insight into personal values can contribute to the learning experience by increasing motivation. The value awareness of the children is the learning objective as value awareness is thought to increase the child’s competence and motivation. One of the options to gain insight into contextual values is Socratic Questioning. These questions are created to further the learning experience and are expected to increase value awareness. This thesis investigates how a social robot can assist children in attaining learning goals, reflecting on these goals, and exploring the decisions and values underlying their learning experiences related to neurodiversity. The effectiveness of the robot in achieving learning objectives, encouraging reflection on decisions and gaining insight into personal moral values is shown. To gain insight into personal values, two methods of Socratic Questioning have been compared; Clarification and Implications & Consequences. The participating children of this thesis play an educational game where moral dilemmas are presented. A social robot assists in the game and asks the Socratic Questions. The explanation to the Clarification has proven to be the most effective method out of the two to gain insight into moral values. Moreover, the interaction effect between question type and time shows that the impact of the question type on insight into values varies with the time a question is asked. Qualitative observations show the participants’ anticipation of the robot’s questions and show instances where responses are different from the self-reported value importance. Overall, this research sheds light on the potential of Socratic Questioning and its uses for understanding individual values, calling for further exploration in this area.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background</b>	<b>2</b>
2.1	Definition of concepts . . . . .	2
2.2	Schwartz' Values . . . . .	3
2.2.1	Insight in moral values . . . . .	5
2.3	Robots in the classroom . . . . .	6
2.4	Raising awareness of neurodiversity . . . . .	7
2.5	Research Question . . . . .	9
<b>3</b>	<b>The game</b>	<b>10</b>
3.1	Goals and requirements . . . . .	10
3.2	Design rationale . . . . .	11
3.3	Educational learning goals . . . . .	12
3.3.1	Representation of neurodiversity . . . . .	12
3.4	Script for the game . . . . .	12
<b>4</b>	<b>Robot design</b>	<b>14</b>
4.1	Reflection on choices . . . . .	14
4.2	Enhance retainment of learning goals . . . . .	16
4.3	Goals of the interaction . . . . .	16
4.4	Robot control . . . . .	16
4.5	Script for the robot . . . . .	16
<b>5</b>	<b>Evaluation Method</b>	<b>18</b>
5.1	Participants . . . . .	19
5.2	Experimental design . . . . .	20
5.3	Assignments for participants . . . . .	20
5.4	Variables . . . . .	20
5.4.1	Pre-session . . . . .	20
5.4.2	During session . . . . .	21
5.4.3	Post-session . . . . .	23
5.5	Material and Experimental setup . . . . .	23
5.6	Procedure . . . . .	26
5.7	Statistical tests . . . . .	27
<b>6</b>	<b>Evaluation Results</b>	<b>27</b>
6.1	Failed data collections . . . . .	27
6.2	Hypothesis 1 . . . . .	28
6.3	Hypothesis 2 . . . . .	30
6.4	Hypothesis 3 . . . . .	32
6.5	Qualitative Analysis . . . . .	36
<b>7</b>	<b>Discussion</b>	<b>46</b>
7.1	Implications . . . . .	47
7.2	Limitations . . . . .	47
7.3	Further Research . . . . .	49
7.3.1	Other types of Socratic Questioning . . . . .	49
7.3.2	Better way of rating answers . . . . .	49

7.3.3	Personality with a bigger sample size . . . . .	49
7.3.4	Add additional types of neurodiversity . . . . .	50
7.3.5	Test the true influence of a game on the motivation . . . . .	50
7.4	Recommendations . . . . .	50
7.4.1	Explanation at the choice question . . . . .	50
<b>8</b>	<b>Conclusions</b>	<b>50</b>
	<b>References</b>	<b>51</b>
<b>A</b>	<b>Questionnaires</b>	<b>57</b>
A.1	Pre-session . . . . .	57
A.2	Post-session . . . . .	63
<b>B</b>	<b>Scripts</b>	<b>66</b>
B.1	The game . . . . .	66
B.1.1	Situation 1 . . . . .	66
B.1.2	Situation 2 . . . . .	68
B.1.3	Situation 3 . . . . .	69
B.1.4	Situation 4 . . . . .	70
B.2	The robot . . . . .	71
B.2.1	Situation 1 . . . . .	71
B.2.2	Situation 2 . . . . .	74
B.2.3	Situation 3 . . . . .	77
B.2.4	Situation 4 . . . . .	80
<b>C</b>	<b>Information form and consent form</b>	<b>84</b>
C.1	Information form parents . . . . .	84
C.2	Information form children . . . . .	88
C.3	Benefits of gaming literature review . . . . .	90
<b>D</b>	<b>Game images</b>	<b>91</b>

# 1 Introduction

In the Netherlands, approximately 1.371.000 children attend primary school [1]. When a child is in primary school, they learn a lot about very important aspects; they learn about standard subjects like mathematics and grammar but they also learn how to act among their peers. Teachers spend time teaching social skills and work actively on bullying protocols in the classroom. Eight per cent of the Groep 8 (the last grade of primary school, ages 11-12) has been bullied [2] which shows the importance of improving social skills and preventing bullying.

An important factor in learning social skills is being aware of one's moral values. When a curriculum is created that only focuses on raising the test scores of basic courses like language and mathematics, it is unlikely to be effective unless virtue is incorporated, valued, intentioned, and practised [3]. Awareness of their moral values can help in character growth, decision-making, empathy, and compassion [4]. It could even influence students' motivation [5]. Motivation can impact the student's learning in different ways. If the student gets motivated to achieve academic success, motivation could even influence the student's learning. The student can also gain the insight that benevolence is of importance to them. In that case, the student could get motivated to learn about other children to become aware on how to be more benevolent.

Integral to a student's learning experience are the teachers. A teacher also functions in a lot of different roles. Teachers are mentors, coaches, support, helping hands, role models, mediators and much much more, this indicates the importance of teachers. Unfortunately, in December 2022, there was a reported shortage in primary-schoolteachers of almost ten per cent in the Netherlands [6]. Due to this shortage, teachers are often overworked and have to take in more children in their classes. In addition, the number of children in a classroom has increased for a couple of years. In 2020, this number was 22.9 children per classroom [7]. The combination of these two factors often leads to less one-on-one time per child.

Due to this shortage and busy schedules of teachers, teachers often do not have the time to cover all topics. The main focus lies on the core cognitive skills like language, writing and mathematics and less on social skills. Consequently, there is hardly time for lessons on neurodiversity and how this affects behaviours in and outside the class. Neurodiversity is an umbrella term for all-natural variations in neurocognitive functioning among individuals. When this topic is handled in a classroom, it can contribute to improving one's social skills and self-awareness. The topic of neurodiversity requires the teacher to have a lot of knowledge and time to cover it and due to the lack of teachers, the time often is not there to educate themselves or the children on this topic.

There is quite a range of conditions that fall under the umbrella term of neurodiversity; some examples are Down Syndrome, Dyslexia, Attention Deficit Hyperactive Disorder, Autism Spectrum Disorder, Epilepsy and even mental health conditions can fall under the term. As fifteen to twenty per cent of people are thought to be neurodivergent [8], it is very likely that there is at least one neurodivergent student in the classroom. Therefore, education on neurodiversity is even more important to improve inclusion for these students. Some aspects of neurodivergency can impact the students in the classroom. For example, some people can have more energy which they are unable to expend due to the students sitting inside for a big part of the day. This can bother other students if the student is not able to sit still. In addition, some aspects of neurodiversity interfere with the ability to understand and exhibit social cues. This could lead in the worst case to misunderstandings and fights.

Robots might assist teachers in the class, reduce their workload and provide means to learn about a specific "social" topic like neurodiversity. Quite a few researchers are developing and testing social robots for classrooms with diverse roles like teachers, teacher assistants, tutors, peers, care-receiving

robots, coaches or mentors. The current research is mainly focused on using robots as teacher assistants [9]. Note that these robots are mainly research prototypes. Only a few social robots are sporadically being deployed in classrooms, which are mainly focusing on assisting in core cognitive skills like math, grammar and programming [9], [10]. Recently, robots have been developed to assist with learning social skills. However, these were mainly used in therapeutic sessions, like improving the social skills of children with Autism Spectre Disorder [11].

This research is part of the ePartners4All project. The ePartners4All project is focused on developing robots that educate children about mental and physical health and well-being, taking into account individual differences among children. When a robot buddy is used, a child can have more one-on-“one”, personal, learning time. This thesis focuses on the interactive reflections that the robot can provide when a child is playing an educational game on neurodiversity. Such reflections intend to raise children’s awareness of the values at stake in attitudes and behaviours towards neurodiverse children. Socratic questioning is a method for such reflections, providing several types of questions to assess a situation or choice [12]. However, how far this questioning can be applied to children in the age of 8-12 and which type of questioning is most appropriate is unknown [13]. Currently, there is hardly any research or development on such reflective dialogues of social robots for children.

## 2 Background

### 2.1 Definition of concepts

The definition of key concepts is required to ensure a shared understanding of terminology. The concepts discussed are educational games, neurodivergence & neurodiversity, and values.

An educational game is a game that is specifically designed to teach people – primarily children – a certain subject or assist in learning a skill while playing. It is often called edutainment due to the combining of an entertainment source such as video games with an educational process [14]. Although educational games seem to be a quite new phenomenon, they have been around in a technological context since 1970 [15]. Educational games are part of a trend called “Serious Games”. Serious games are games where the game’s primary purpose is not to entertain.

An additional relevant concept to consider is the concept of neurodiversity and neurodivergence. These terms are occasionally used interchangeably but have different definitions. Shah et al. describe Neurodivergence as the following: “Neurodivergent describes individuals whose selective neurocognitive functions/ neurodevelopmental differences fall outside prevalent societal norms. They do not necessarily have a neurodevelopmental disorder” [16] (p. 579). They describe neurodiversity as: “Neurodiversity is the statistical normal range of function in a population at a particular age. Neurodiversity is a characteristic of the whole population, not a specific individual.” [16]

The most common types of neurodivergence are Attention Deficit (Hyperactivity) Disorder, Dyslexia and Autism Spectrum Disorder. However, there are more types that are less familiar like Tourette’s, Dyspraxia and Dyscalculia. When talking about a neurodiverse community, it can consist of a combination of neurotypical and neurodivergent people. Using the proper term to refer to neurodivergence or neurodiversity is quite important, Legault et al. [17] even consider the term “neurodivergence” as a term of exclusion while they consider “neurodiversity” as a term of inclusion. When talking about a population, the term neurodiversity is more inclusive. When talking about people whose neurocognitive functions and neurodevelopmental differences fall outside of the societal norms, the term neurodivergent is used in this paper.

The types of neurodivergence that will be targeted in this paper are Attention Deficit (Hyperactivity) Disorder (AD(H)D), Dyslexia and Autism Spectrum Disorder (ASD). These types of neurodivergence are chosen due to them being the three most common types [18]. The American National Institute for Mental Health describes ADHD as: “being marked by an ongoing pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development.” [19]. They describe ASD as “a neurological and developmental disorder that affects how people interact with others, communicate, learn, and behave.” [20] The American National Institute of Neurological Disorders and Stroke describe Dyslexia as “a brain-based type of learning disability that specifically impairs a person’s ability to read. Despite having normal intelligence, these individuals typically read at levels significantly lower than expected.” [21]

Finally, the concept of moral values is of importance. A value is something that is important to an individual. Each person holds multiple values with varying rates of importance. Examples of values are Achievement and Power. As mentioned in section 1, being aware of one’s values could have an influence on raising a student’s motivation [5]. Having a higher motivation to learn impacts the student’s achievement in school [22]. Moral values in the context of this research are considered the list of ten values that Schwartz defined [23]. These values will be elaborated on in subsection 2.2.

## 2.2 Schwartz’ Values

Schwartz has indicated six important features to values [24]; Firstly, a value is a belief. When a value is activated, it becomes connected to emotions and feelings. Secondly, values refer to goals that motivate action. Thirdly, values stay the same even when the scenario changes, this makes values differ from norms and attitudes. Fourthly, values serve as standards or criteria, they guide certain actions and decisions. However, when a decision is made based on values, the person is rarely conscious of their values. Fifthly, values are ordered by importance. Finally, there is a relative importance of multiple values guiding actions. Multiple values can be connected to a single action.

These six features of values were used by Schwartz to define 10 values: Self-direction, Stimulation, Hedonism, Achievement, Power, Security, Conformity, Tradition, Benevolence and Universalism. Schwartz explains these 10 values based on the goal it expresses [23]. In addition, the goals will be linked to situations that can occur in the daily life of children.

### *Self-Direction*

The goal of Self-Direction is independent thought and action like choosing, creating and exploring. It is derived from the need for control and mastery. Children can encounter Self-direction when they are allowed to choose what they are going to work on in the classroom.

### *Stimulation*

The goals for Stimulation are excitement, novelty, and challenge in life. It is derived from the need for variety and stimulation. This can be encountered when children are going on a field trip with the class, this could be considered as an exciting trip where they can encounter a new and challenging situation. Educational games can also be considered as stimulation as they often incorporate excitement, novelty and challenge.

### *Hedonism*

The goals of Hedonism are pleasure or sensuous gratification for oneself. Children can often consider playing and being creative as actions that make them happy.

### *Achievement*

Achievement has as defining goal to achieve personal success through showing competence. This can result in increased self-respect and social recognition. This is closely related to a school setting,



children are given tests where performing well results in the showing of competence.

#### *Power*

Power has multiple defining goals. The first is gaining social status and prestige. The second one is having control or dominance over people and resources. Both power and achievement have social esteem as a relevant factor. However, achievement is closely connected to the active demonstration of successful performance while power emphasizes attaining or preserving a position of power. Power is a concept that might not be obviously prevalent in the daily life of a child. However, children are confronted by the concept of power when they are working on a group project, as one child needs to take the lead to ensure that the group project will be brought to a successful end. In addition, children are often taking the lead in games and sports, this is also a position of power that they are confronted with. Finally, when a child starts bullying, they ensure power over another child in a negative manner.

#### *Security*

The defining goal of security is safety, harmony and stability. This can be in society, relationships or of themselves. A child is often not aware of the concept of security as adults are often there to keep the child safe. In the life of a child, security is often prevalent in ensuring stability in relationships with friends, and by ensuring that they will not get harmed. An example of where security is impacted is when a child gets bullied; the safety of themselves and of the relationship with the bully gets impacted.

#### *Conformity*

In conformity, a defining goal is the restraint of actions, inclinations and impulses that would upset or harm others or violate expectations and norms. In conformity, obedience, politeness and respect are of importance. Conformity often occurs in a child's life as they are required to listen to the adults in their lives. They are required to be obedient and polite and show respect to their elders. Children are also conforming to other children. When a topic becomes popular in the classroom, like TikTok, children are often following the opinions of their classmates to not upset others in their class.

#### *Tradition*

In tradition, respect and acceptance of a culture's customs and ideas are important. Tradition is usually connected to more abstract ideas and customs. Tradition in a child's life often is presented by the adults in their life. Certain traditions, like putting up a Christmas tree, are introduced to the child. As children are quite young, they have only recently been introduced to, and fostered awareness for traditions.

#### *Benevolence*

A defining goal for benevolence is preserving and enhancing the welfare of others. Benevolence values provide an internalized motivational base for behaviour. In the daily life of a child, benevolence is very important. Children are often helping each other which could increase the welfare of the assisted child as a challenge or burden might be reduced..

#### *Universalism*

In Universalism, the defining goal is the welfare of all people and nature. It is important to understand welfare, appreciate it, tolerate it and protect it. This is quite an abstract value, closely related to inner harmony and a spiritual life. Examples of universalism are; unity with nature, protecting the environment, and the world at peace. In this context, we relate it to the benefit of a bigger group, like the entire class. Children are often confronted with decisions that could benefit their entire classroom, like how they should behave to ensure that the entire class does not get in trouble. Most schools introduce bullying prevention protocols which allow for the enhancement of the welfare of the entire classroom.

### 2.2.1 Insight in moral values

To enhance the insight into moral values, some methods have been created like the Schwartz Value Survey (SVS)[25], Portrait Value Questionnaire (PVQ) and Picture-Based Value Survey for Children (PBVS-C) [26].

Schwartz Value Survey consists of a list of 56 items [25]. These items in the survey are all represented by one of eleven values. One value was later removed by Schwartz. To find the underlying values, participants in that survey are asked to rate the importance of these items to them on a 9-point Likert scale where numbers are used. Their answers are mapped to the corresponding values. This system of finding values induced some trouble in certain population groups. Doring notes that the survey demands a high level of abstract thought which is why the SVS is not suitable for people who are not educated in context-free thinking [27]. In addition, the survey is considered a very long survey as it has 56 items. People who are not as motivated or have difficulties concentrating for a long time might encounter difficulties filling in such a long survey.

To solve this issue of the survey not being suitable for people not educated in context-free thinking, the Portrait Value Questionnaire (PVQ) has been created by Schwartz[28]. The PVQ is more contextualized and concrete. Doring evaluated the PVQ with children [27]. Doring noted that the terms used in PVQ often required additional explanation by the researchers. They also noted that participants seemed to miss items in the questionnaire, most likely due to the reduced attention span with proceeding time. They indicated that the children often found the PVQ a challenging instrument.

This led Doring et al. to create the Picture-Based Value Survey for Children (PBVS-C) [26]. This Survey is considered the first self-report instrument for analyzing young children's values using Schwartz's ten values. The survey uses pictures with a small caption to explain the values to the participants. The participants are expected to rank each of the values from "not at all important" to "very important". The results of this study indicate that the PBVS-C is considered a good tool to find the values of children. The limit to this PBVS-C is that the context of a situation is not considered, this shows a simplified version of reality. However, context is of great impact to decision-making. Yudkin et al. showed that people select different moral values according to the situation, and how the mere presence of others can affect moral thinking [29]

#### 2.2.1.1 Socratic Questioning

Another method to gain insight into moral values is to reflect on values with Socratic Questioning. Socratic Questioning can be used with more context to ensure that the representation of reality is less simplified, allowing for a more comprehensive and nuanced understanding.

In Socratic questioning, questions are set up in a specific way to probe to a deep level of thinking. Socratic questioning is a method named after Socrates as Socrates used to ask his students questions to enhance learning. According to a literature review by Carey and Mullan, Socratic questioning has a variety of different definitions and uses [30]. One of the examples of the definition of Socratic Questioning they give is by Overholser [31]; "The Socratic method of interviewing encourages the client to contemplate, evaluate, and synthesize diverse sources of information, most of which were already available to the client." The Socratic conversation of which Socratic Questioning is part of, was created to find eternal values [32]. Socratic Questioning is also commonly used in pedagogy to probe student thinking. The method of Socratic Questioning can enhance both the students' speaking skills and their critical thinking [12]. This method of questioning has also been evaluated in the context of moral reasoning. Torabizadeh et al. have evaluated the impact of Socratic Questioning on the moral reasoning of nursing students [33]. They have shown that Socratic questioning was considered more

effective than lecturing in teaching ethics in improving the subjects' moral reasoning skills.

Six different types of Socratic Questioning have been created. These types all address various aspects of thinking. Each type specifically encourages critical thinking, uncovers underlying assumptions, promotes self-discovery, and creates insightful conversations. Manurung has evaluated the six types and ordered them in a range based on cognitive difficulty. Successively, the types involve more in-depth thinking, increasing the difficulty to answer [12]. So what type of question is best, is dependent on the goal and the cognitive skills of the questioned person.

The easiest type is "Questions for Clarification". These questions aim to get a better or further defined answer. Examples of questions are the following: "What do you mean when you say X", "Could you explain that point further" and "Can you provide an example?" [34]. These questions are often used in education to see whether a student understands a concept and if they can explain their reasoning. These questions mainly aim to let students think further and deeper about the subject [12].

The next type is "Questions that Probe Assumptions". These questions aim to challenge assumptions that they have been made. Examples of these questions are the following: "What assumptions are we making here?", "Are you saying that...?"[34]. Probing assumption questions make students think about presuppositions and unquestioned beliefs on which they are finding their argument [12].

The third type of questioning is "Questions that Probe Reason and Evidence", which means that the questions are formulated to let the students provide evidence and reason as a basis for their arguments. Examples are: "Can you provide an example that supports what you are saying?", "Can we validate that evidence? Do we have all the information we need?" [34]. These types of questions support students to dig into reasoning rather than assuming it was given while people often use weak supports for their arguments [12].

The fourth type is "Questioning about Viewpoints and Perspective". Questions like this often aim to find alternative viewpoints. Examples are: "Are there alternative viewpoints?", "How could someone else respond, and why?" [34]. These questions criticize the position and show the existence of equally valid viewpoints. [12].

The fifth type is "Questions that probe Implications & Consequences". These types of questions aim to let the students explore the Implications & Consequences of situations. These questions can include: "How would this affect someone?", "What are the long-term implications of this?"[34].

The most difficult question type is: "Questions about Questions". The students explore why a certain question is asked and how it could be relevant to their learning. Example questions are: "What do you think was important about that question?", "What would have been a better question to ask?" [34].

In this paper, the method of Socratic Questioning will be used to facilitate children's insight into their own values while also providing researchers with valuable insights into the values held by the children. Socratic Questioning has not previously been used in the context of moral values in children but it has been used to enhance moral reasoning of nursing students [33].

### **2.3 Robots in the classroom**

Robots have been used in classrooms in quite a few studies. A literature study done by Rosanda and Istenic Starcic, researched nineteen papers and identified that the robot was used as a teacher most often, the robot performed the role of teacher in 29% of the cases [9]. In most of the research considered in the literature study, the robot was focused on delivering learning materials and not on individual

teaching, which includes delivering feedback and tailoring learning activities to the individual. To see if a robot could be considered as an adequate teacher, Edwards et al. tried to see if a robot as a teacher would be deemed credible [35]. They concluded that adult students deemed a robot as a teacher credible.

Robots are most often used as teachers. However, Tanaka and Matsuzoe used the robot as a peer as well [36]. In their research, they used a robot to teach children about caregiving. For their test, they performed lessons given by a human teacher with the participant children. In some cases, the robot was used by the teacher as a peer to the child participant, the robot often made mistakes where the teacher would correct them. Afterwards, the child was allowed to teach the robot when it would make mistakes. They reported that the robot contributed to the enhancement of spontaneous learning. They also concluded that using a robot resulted in a significant difference in the questions that were answered correctly, with the robot, more questions were answered correctly.

Another recent study is the PAL project. In this project, a robot was used to improve the health literacy, attitude and motivation of children with diabetes type 1 [37]. It was shown that the combination of a social robot and an educational game worked well. The focus of the project was on goal-setting, feedback and explanation. In one of the experiments within the PAL project, the use of a personal robot was compared with a neutral robot, showing that the personal robot was more effective in increasing health literacy on diabetes. Another study by Peters et al. within the PAL project focused on goal setting [38]. They analyzed the way the robot could assist with learning goals and suggested guidelines for a framework where goal-setting is supported. They suggested four guidelines related to the user interface and authoring tool of a goal-setting framework.

Within the ePartners4all project, van den Berg did a study to educate children to improve their dietary decisions[39]. Van den Berg showed that children were motivated to learn with a robot and had a good learning performance with the robot. This shows that learning with a robot is a feasible option.

## 2.4 Raising awareness of neurodiversity

Finally, we take a look into related work in raising awareness about neurodiversity.

Games have been created to assist neurodivergent students in some aspects of their daily life. Boyd et al. [40] created three different Virtual Reality games that help people with Autism Spectrum Disorder in their daily lives. The first game; Bob's Fish Shop, aims at practising conversational skills, nonverbal communication and joint attention skills. Their second game; VirtualBlox, is designed to exercise fine and gross motor skills. Their last game; vrSocial, aims to support nonverbal communication by visualizing proximity, volume and duration of talking. After creating these three games, their conclusion is that Virtual Reality is a great method to attend to the needs of neurodivergent people as one is able to adapt the environment to the tailored needs of their user. However, these games do not attempt to simulate neurodivergence nor do they raise awareness. They mainly focus on assisting neurodivergent users in their daily lives.

Santhanam proposed a social space in a video gaming context where the aim is to connect autistic and non-autistic people and to raise awareness about autism [41]. They suggest connecting autistic and non-autistic students based on their interests. The non-autistic students in this experience were following a service learning class on autism. This method of combining autistic and non-autistic students showed that the non-autistic students gained an awareness and appreciation of the differences in communication. The examples that are shown are that they gained an understanding of long silences, lack of eye contact and the lack of small talk. The experiment was done using video games

that already exist. They have chosen to use teamwork video games where collaboration is usually required. The games were not developed specifically to raise awareness or to educate, rather they were created to have a fun time and to collaborate with your peers.

Some scientists, artists and many others have tried to simulate the experience of being a neurodivergent person. Young is an artist that has created a virtual environment that tries to raise awareness on the daily difficulties of people due to their dyslexia [42]. López-Cara ral et al.[43] have simulated the experience of having Autism Spectrum Disorder (more information on ASD can be found in subsection 2.1). They have attempted to simulate the over-sensory sensation that people with ASD can experience. The virtual reality experience is enhanced by biofeedback using the users' physiological signals. Unfortunately, the authors of the paper have only implemented a virtual reality simulation in their paper, however, no evaluations have been mentioned to be performed. Evaluations that might be useful for this simulation would have been both on the effectiveness of raising awareness and to test how near it is to the genuine experience of a person with ASD. In addition, simulation reduces the complexity of the human body into a few separate elements [44], reducing the realism of the simulation.

In the case of Dyslexia, some simulations have also been created. An example of a simulation is Dyslexia simulator, on the website of Harvard. On this website, the letters keep shifting positions within the word. Another website is Expertisecentrum Dyslexie, on this (Dutch) website, there is a story about children going to the funfair. There are six different reading problems that all are represented in a specific subset of the story. After every subset, a question is asked about the story. Some of the problems represented are: interchanging the b,d,p and q, reading letter by letter and changing the order of the words within a sentence. Wadlington et al. did a simulation with four stations where every station represented a challenging task related to dyslexia [45]. As the participants of the study had already obtained quite an extensive background knowledge in their classes, the goal was not to inform the participants but to raise awareness of the needs of students with dyslexia. Wadlington concluded that their dyslexia simulation is a powerful learning experience that can enable teachers to better help all students, especially students with learning difficulties. In addition, a font has been created by graphic designer Daniel Britton. Stark et al. have conducted an experiment to test if this font would simulate the experience of reading a text while having dyslexia [46]. They concluded that 84.5% of the typical readers would fall below the dyslexia group's median while reading this font compared to people with dyslexia reading the same text in Times New Roman font. They concluded that there is no evidence for the font empirically simulating the difficulty of dyslexia as both behavioural reading performance and the visual sampling strategy were not matched. However, increased frustration with reading is highly likely.

We have now considered both video games in the context of neurodivergence and methods to raise awareness of neurodiversity but there is a very limited amount of video games that raise awareness of neurodiversity. Meinen reviewed two video games that were highlighted by the neurodivergent and disabled community on Xbox [47]. The games reviewed were Celeste and Unravel. The games mentioned are both considered metaphors for disabilities and neurodivergence and do not clearly mention the character that the player controls being neurodivergent or disabled. Unfortunately, these games do not attempt to educate the player on neurodiversity or to specifically raise awareness. No attempts seem to have been made to educate and raise awareness about neurodivergence and neurodiversity by using an educational game.

Finally, we take a look into using robots to raise awareness of neurodiversity. Robots have been used in the context of neurodiversity. There has been a lot of research on assisting children with autism in learning social cues. The robots were successful in evoking prosocial behaviours such as joint attention and imitation [48]. The role of the robot in these scenarios is most often the role of a leader, showing the social behaviour and leading the interaction. It is also used as a toy, reacting to the child with

autism and mediating between the child and their peers.

## 2.5 Research Question

The objective is to develop dialogue between the child and robot that increases the child’s value awareness. We will use neurodiversity as a case study. The general research question is:

RQ: How can a social robot support a child in gaining insight into their values behind decisions or choices while learning about neurodiversity?

The method of Socratic Questioning is used for reflective dialogues but has not been studied for children and not within the field of child-robot interactions. Socratic Questioning is used in this thesis to gain insight into values. Insight into values could lead to more motivation [5] and more motivation could lead to better learning. We expect that, due to learning, a consecutive Socratic Question provides better value-based argumentations. In addition, two types of Socratic Questions are of specific relevance (see 2.2.1.1), with a different difficulty-trade-off. So, we will test and compare these types.

Clarification is the easiest type of Socratic Questioning[12], and thus less complex. The reduced complexity compared to “Implications & Consequences” means that this type could result in easier-to-answer questions and thus a better argumentation quality. Leading to the hypothesis:

*H1a: The “Clarification” type of Socratic Questioning by the robot will lead to a better (value-based) argumentation quality of the child than the Socratic Questioning type of “Implications & Consequences”.*

As the questions are repeated over time, there might be a learning effect that occurs.

*H1b: A recurrence of a Socratic Question will lead to a better (value-based) argumentation quality than the previous Socratic Questions.*

Due to the question types being different in complexity, the learning effect may also be different for Clarification and Implications & Consequences. This could also result in an interaction between the recurrence of the question and the type of question. This leads to the following hypothesis:

*H2: There exists an interaction effect between the recurrence of the questions and the type of question that results in a significant impact on the quality of argumentation of the participant*

In addition to the formulation of the questions itself. Differences in the individual might also impact the quality of argumentation. For example, extraversion and openness could help in providing more detailed argumentations on the request of the robot.

The Big 5 personality traits are gathered for every participant. They represent the following aspects: Extraversion, Neuroticism, Openness, Conscientiousness and Agreeableness. These personality traits might have an impact on the insight of the participant in their values. This leads to the following hypothesis:

*H3: The personality traits of a participant have an impact on the quality of argumentation of the participant*

## 3 The game

The game is an educational learning tool that focuses on teaching the child about ADHD, ASD and Dyslexia and on increasing insight in the values that underlie dealing with neurodiverse people. The game has some requirements to be an educational learning tool and to work together with the robot. In this chapter, the goals of the game will be described in subsection 3.1. The design rationale in subsection 3.2 in combination with the educational learning tools subsection 3.3 will then show how these goals got tackled and handled in the game. Finally, the script of the game will be shown in subsection 3.4.

### 3.1 Goals and requirements

Five core requirements have been defined for the game: Informative, Representation, Dilemma, Navigation, Robot Teamwork. These core requirements ensure that the game can properly work together with the robot and learn the child about neurodiversity.

**Informative:** The initial main goal of the game is that it should be informative about neurodiversity. Media about neurodiversity often only portray negative aspects or enhance stereotypes about neurodiversity. This game should try to not only show negative but also positive aspects and should not enhance the common stereotypes. In addition, some media portray neurodivergent people as radically different from neurotypical people. That stereotype should not be enhanced. Only ADHD, ASD and Dyslexia are currently represented, it should be easy to adjust the game for more types of neurodiversity.

**Representation:** The world as we know it does not consist of only neurodivergent or neurotypical people but there is a combination in every classroom. When the game only has neurodivergent people represented, it is not a great representation of the world. Therefore, the game should show both neurodivergent and neurotypical people.

**Dilemma:** To allow for the exploration of moral values within context, the game should give a dilemma. Within the ePartners4all project, a study is done that gave the participants dilemmas with some pictures and asked about which decision they would make and why. A similar dilemma should be set up by the game to allow for this exploration of values. To ensure that the children are making decisions as themselves, the game should give the feeling that the children are playing as themselves.

**Navigation:** The game should be played by children on a laptop. To ensure that they are able to play the game, it would need to feel natural how the controls work and actions (like interacting or opening a door), should be explained by the game. The game should not be too difficult as a difficult game raises the barrier to playing the game, it should be simple, yet intriguing. It should also not feel like a mandatory learning tool but as a game to enhance the children's will to continue playing the game.

**Robot teamwork:** Finally, the game will be played by a child and a robot. The game should thus allow and give space for the robot to converse with the child, as an additional learning tool. The robot asks the child to read the learning goal, thus the game should allow plenty of time for the robot to get the cue to talk, and for the robot to finish the question. After each level, the robot asks the child questions to gain insight into the values behind the decision. To ensure that the child is not distracted, the game should pause until the robot is done talking.

## 3.2 Design rationale

The game has a specific set of features that are considered design choices to fulfil the goals and requirements.

Initially, the child should be able to start with creating their character. Creating a character to represent the child is necessary to ensure that the child feels represented. The child must feel that the decisions that they have to make are not decisions that the character makes but that they as a person make. Having a character also increases the feeling of autonomy for the child [49]. In addition, Trepte et al. discovered that games that include character creation, have a higher rate of enjoyment [50].

In the game should be learning goals. This will make the difference between having a regular game and an educational game. These learning goals will teach the students about certain aspects of neurodiversity. The learning goals will be shown as text on a screen and repeated in the interaction with NPCs. The learning goals are further explained in subsection 3.3

Next, the children should be able to make a decision that reflects their values. They are given a dilemma. For example, they know that Peter will be very good to pick and play music during a party but the child has just learned that Peter does not like busy settings but is willing to do it as a favour. Easily being overstimulated is a trait that people with ASD often, but not always, inhabit. They also know Jan is really enthusiastic to be the DJ, but are a bit worse than Peter. Their dilemma is between having the perfect person for the job or choosing someone to make them happy.

The other people, like Peter and Jan, are Non-playable-characters. Having these NPCs will allow for simulated interactions and the ability to put learning goals in their conversations. These NPCs represent neurodiverse children and they allow for showing certain aspects of neurodiversity. There are also other students in the classroom that the child does not necessarily need to interact with. These children represent neurotypical people. The conversation that they have if interacted with, will give the child more information about the task the child has to do.

To make a certain decision, children should be able to walk towards other (non-playable) children as that allows for some time to think about their decision. It will need them to walk towards a person instead of just clicking the screen left or right. This will allow the child to have a bit more time to contemplate their decision and they take a more active action. Having the possibility to walk towards a character will also make the interaction feel more life-like.

The inside of the school that is designed in the game is modelled to an average Dutch primary school (“Basisschool”): The tables are set in groups of four, there is a screen with the teacher’s desk next to it, there is art on the wall, and there is a teacher’s room with a small kitchen. This design choice is to enhance the feeling of the child that they are in a primary school. Which will enhance the feeling of autonomy as they are familiar with the layout.

To control the character in the game, the child can use both the arrow keys and WASD to walk around the screen. On the top bar, the task that the child should do is shown. On the bottom bar, the learning goal is shown.

When the child gets close to a door, a little text box pops up on the screen, instructing them to press Enter to enter the door. On the top bar of the screen, left of the task that the child should do, a little text box is shown that tells the child to press Z to interact with another person in the game.

At the end of each level, a pink screen will appear. The text on the screen tells the child to listen to



the robot and to not click on Done before the robot says so. This screen gives the robot time to talk to the child and ensures that the child is not distracted by the next level yet.

### 3.3 Educational learning goals

The difference between a regular game and the educational game is that an educational game teaches specific learning goals to the player of the game. The game consists of four levels. There are two learning goals in level one, and one each in levels two, three and four. This results in five learning goals. No more learning goals are added due to the difficulty of learning more than five learning goals in half an hour.

An essential factor in the learning goals is that they should not enforce stereotypes. Not all people with neurodiversity exhibit the same traits. To ensure that this stereotyping does not occur, the learning goals in the game are more generalized.

The learning goals are the following:

1. Some children can handle busy environments a bit worse than others, they might get overstimulated
2. Some children have a bit more energy to expend than others
3. Some children are more hesitant to ask something to the teacher than other children
4. Some children have more difficulty understanding jokes than others
5. Some children have to put more effort into reading and writing without mistakes than others.

The first, third and fourth learning goals are aspects that often, but not always, occur in people with ASD. The second learning goal is an aspect that often occurs in people with ADHD. The fifth learning goal is an aspect that often occurs in people with Dyslexia.

#### 3.3.1 Representation of neurodiversity

The Non-playable characters that the child is talking to will show the traits mentioned in the learning goals accompanying that level. As referenced in the definition of the learning goals, the traits do not enforce stereotypes. The problem that the children that play the game encounter are in line with the learning goals. In this manner, the learning goals are repeated and shown in context to better help the understanding and remembering of these goals.

### 3.4 Script for the game

In this section, the script of the game will be explained. The situations will be explained, and the matching learning goal and the task for the child. The script for the first level is shown for illustrative purposes on how a level looks like. All of the levels in English and Dutch can be found in subsection B.1. Images from the game can be found in Appendix D.

*Italics:* Things that the characters in the game say

**Bold:** Name of NPC who says something

**Blue:** childname, to be filled in with name of child

Underlined: Button/Option to be clicked by child

**Green:** The action that the child does

**Orange:** Text of learning goal in game (bottom of screen)

**Pink:** Text in the game in the middle of the screen

#### Level 1

In the first situation, there are 2 characters: Jan and Peter. Peter has some aspects of ASD. Jan has

some aspects of ADHD.

There are two accompanying learning goals: The one accompanying Peter is: Some people can handle busy situations a bit worse than others.

The task for the child is to choose a DJ to play music. They can choose between Peter and Jan.

The learning goal accompanying Jan is: Some people have a bit more energy that they need to expend than others

## Script

The child starts in the classroom with the teacher. Peter is in the auditorium, Jan is on the basketball field.

**Teacher Evelien:** *“Hello **childname**, I need some help planning the party, can you help me by any chance? I need help choosing a DJ, I heard that **Peter** is a super good DJ but maybe he doesn’t want to be a DJ. I’ve also heard that **Jan** might want to be a DJ. Could you choose who will be the DJ? **Peter** is in the auditorium and **Jan** is on the basketball court. Come back to me when you have talked to them.”*

Option button: *“Okay!”*

**Action:** The child walks in the school

**Text during action:** Some children can handle busy environments a bit worse than others, they might get overstimulated.

**Action:** While the child walks outside

**Text during action:** Some children have a bit more energy to expend than others.

Conversation with **Peter:**

**childname:** *“Hi **Peter**, I have been instructed by the teacher to find a DJ. The teacher said that you are really good.”*

**Peter:** *“Hi **childname**. I can choose music pretty well, but at a party, it is often too busy for me, but I can be a DJ.”*

Conversation with **Jan:**

**childname:** *“Hi **Jan**, I have been instructed by the teacher to find a DJ. The teacher said that you might want to do it.”*

**Jan:** *“Yesssss! I would love to be a DJ! I can also do the decorations but I would really prefer to be a DJ!”*

**Action:** The child returns to the teacher

**Teacher Evelien:** *Who do you want to make the DJ?*

Option button 1: **Peter**

Option button 2: **Jan**

**Text on the screen:** The robot will now ask you a few questions, click done when the robot says you can press done.

Option button: Done

## Level 2

A student will approach the child and tell them that they have difficulty with grammar but are afraid to ask the teacher because they feel that their question is stupid. They ask the child if they could ask the question to the teacher for them.

The task for the child is to choose whether or not to ask the question to the teacher as if it was their

own question or to tell that it was a question from the other student.

The accompanying learning goal is: Some children are more hesitant to ask something to the teacher than other children.

### **Level 3**

In this scenario, two students are on the playground. One of the kids makes a joke, and the other kid does not understand the joke and gets angry. There is implied that a fight might start and the child needs to choose what to do.

The decision for the child is to choose whether to go to the teacher so they can intervene or to jump in between the two angry children themselves.

The learning goal is: Some children have more difficulty understanding jokes than other children.

### **Level 4**

The teacher is making groups for a reading and writing task. The child is grouped with Mike. The child knows that Mike has difficulty reading and writing as fast as the child. The child can decide to do most of the work themselves or to give Mike an equal share of work. The child needs to choose between doing most of the work themselves or spread the tasks equally but possibly getting a worse grade or the task taking longer.

The learning goal is: Some children have more difficulty reading and writing than others.

## **4 Robot design**

### **4.1 Reflection on choices**

To compare getting insight into the values of students, we would like to work out a conversational reflection method that provides insight in the values underlying decisions (i.e. the choices on value-related dilemmas). We first need to identify what methods can be expected to provide such insight and then we need to test if this expectation proves to be correct. The suggested method in this thesis is to use two of the types of Socratic Questioning.

Socratic Questioning consists of six different types of questioning: Questions for Clarification, Questions that Probe Assumptions, Questions that Probe Reason & Evidence, Questions about Viewpoints & Perspective, Questions that probe Implications & Consequences and Questions about Questions.

Although Socrates designed the Socratic Questioning to find eternal values [32], not all of the methods are probably equally useful for children. The first method of Clarification, where questions are asked that let students think more about what they were asking and thinking about [12], seems to be quite useful for getting values as it lets the student gain insight into why a decision has been made. The second method, Probing Assumptions, which uses questions to probe about unquestioned beliefs on which they are founding their arguments, is more difficult to use as children might not understand the concept of assumptions at a young age. The method of Probing for Reason & Evidence, that uses questions to dig into reasoning, is more difficult due to the concept of the test; children are given a choice which is based on learning goals, and it would be rather easy to repeat the learning goals as evidence of a choice which would not accurately reflect the value of the learner. The fourth method of Finding Viewpoints & Perspective, searches for an alternative viewpoint in the situation. While this is very useful to show alternative methods, it shows the viewpoint and values from a different perspective and not from the learner. The fifth method asks for the Implications & Consequences of a situation. This method is useful in extracting the values of the child as they show what the implication of their choice was and how it relates to the future. If they have thought about the consequences in making their choices, it will be related to their values. The type of Question about the Question, uses

Socratic Questioning type	Description	Applicability value statements	Complexity (for children)
Clarification	Clarify thinking and explore origin of thinking	Useful	Easy
Probe Assumptions	Probe about unquestioned beliefs on which arguments are founded	Not useful	Difficult
Probe Reason & Evidence	Dig into reasoning	Medium Useful	Medium
Viewpoints & Perspective	Search for alternative viewpoint in the situation	Not useful	Medium
Implications & Consequences	Show implications of certain decisions	Useful	Medium
Questions	Explore why certain questions are asked	Not useful	Difficult

Table 1: Types of Socratic Questioning, table includes the description of the type, the applicability to analyze value statements and the applicability for children in regards to complexity.

questions to show why certain questions are asked. This method is too abstract to find values, it will show if the children understand why certain questions have been asked and to find the learning goals. Table 1 shows the six types of questions, their explanation and their applicability in this experiment. The two most applicable options are “Questions for Clarification”, and “Questions that probe Implications & Consequences”. This is also an interesting combination as the first option is deemed the easiest and the fifth is one of the most difficult options. This will allow us to compare if the difficulty of the question also has an impact.

In “Questions for Clarification”, one of the most commonly used keywords in these questions is the word Why. To address why questions, we first need to identify how the question should be formulated. It is identified by Bateman, that task-related and emotional-related questions can be handled in a similar way which blurs the line between educational and emotional support in early-child education [51]. Many researchers have emphasized that it is better to focus on reflection than on interrogation, this also shows empathy and understanding. Cameron reports that reflective skills provide a solid basis for establishing rapport, getting the interview underway and encouraging children to tell their story [52]. Why questions are considered as more difficult to answer. Wilson and Powell report that “What”, “Where” and “Who” questions are possible to answer from the age of three [53]. “Why” and “How” are more difficult to due to these questions being more abstract. The “What”, “Where”, “Who” questions being more concrete. Children may only be able to consistently answer “Why” and “How” questions when they are five or six years old. In addition, these questions appear more accusatory and imply blame. Wilson and Powell recommend being careful when asking these questions to children and not implying blame. Intonation and emotion impact the meaning of a statement. For example, a question like: Why did you do this? Can mean something different depending on where the intonation lies. If the intonation lies on “you”, the question is immediately considered as more accusatory then when the intonation lies on “Why”. The advantage of using a robot is that there is no emotional and intonation effect. This will ensure that the questions are less likely to be accusatory or imply blame. In addition, the implication of the choice is not shown. This ensures that there are no right and wrong decisions and thus no blame could be applied.

In the other method, “Questions that probe Implications & Consequences”, common questions ask about how a certain situation can affect another person and how certain situations will result. To ask these types of questions to children, the concept of Implications & Consequences of actions need to be clear. Children learn consequences between the ages of 6 and 13 [54]. The children in this experiment will be in the middle of this age range. This means that they have a basic but still developing awareness of the consequences of their actions. However, the questions need to be formulated in an understandable way for the children. The words consequences and implications are rather difficult words. Thus the questions need to be formulated age-appropriately. To make the concepts more appropriate for children, easier words are used. For example, “what are the consequences of your action?” would become “what would happen if you do this?”

The benefit of “Questions for Clarification” is that the questions are considered easier to answer than the Implications & Consequences questions. The Implications & Consequences questions are more

difficult which could result in lower quality of argumentation due to the child needing to get familiar with the style of questions. Consequently, there is more to gain with the Implications & Consequences type of answers as they start with lower initial answers.

## 4.2 Enhance retainment of learning goals

As mentioned in section 3, the game includes a set of learning goals about neurodiversity for the students to learn. The robot should enhance the learning experience and help in retaining the learning goals. A previous study by Merel van den Berg showed that children are able to retain learning goals when learned with a robot [39]. The robot helps in two methods. The first method is asking the child to read the learning goal out loud to them when the learning goal appears in the game. If the children are unable to read the learning goal, the robot will tell the learning goal to them. When the child successfully reads the learning goal to the robot, the robot will complement the child on reading the learning goal. The robot also tells the child that they are learning together and asks them to remember the learning goal. At the end of the level, the robot asks if the child remembers the learning goal(s) of that level. If the child remembers the learning goals and tells them to the robot, the robot will praise them. If the child does not remember the learning goals or remembers them wrong, the robot will tell them that they are doing great and will repeat the learning goals to them.

## 4.3 Goals of the interaction

The robot should assist the child in achieving the learning goals  
The robot should assist in reiterating the learning goals after the game is done  
The robot should let the child make its own decisions, without judgement  
The robot should ask why certain decisions have been made

## 4.4 Robot control

To reach the goals of the interaction, there are two control goals: the robot needs to accurately respond to the child and the robot needs to be aware of when certain triggers in the game occur. To fulfil the first of these two requirements, the robot needs to understand what Dutch children are saying. Unfortunately, the speech software is not yet at the level where it would accurately understand Dutch children. To mitigate this issue, the researcher listens to the robot and selects how the robot should respond. The participant is unaware that the researcher directly controls the robot. This method of control is called Wizard of Oz. To ensure a speedy response, a script has been created that allows the researcher to choose between buttons to evoke speech from the robot. The robot asks a question to the participant based on the script. After this question has been asked, buttons or a text field appear on the researcher's screen. These buttons correspond to all of the possible response options from the robot. The option of repeating the question in a slightly different manner also corresponded to a button on the researcher's screen. When the child showed misunderstanding or could not be heard well, this button was pressed. For example, if the question was: "What would happen if you jumped in between the fight?", an alternative question would be: "What happens if you do this?", This slight alteration could ensure that the child would know what to answer.

## 4.5 Script for the robot

Per situation mentioned in subsection 3.4 the robot will ask 4 questions. The questions will be either Questions for Clarification or Questions that probe Implications & Consequences as mentioned in subsection 4.1. When a child answers with "I don't know", the robot will ask a reformulation of the question. If a child answers "I don't know" to a Clarification question, the robot will ask another

Clarification question until there either is an answer, or the researcher feels that the child will not be able to provide a meaningful answer. The same process occurs for Implications & Consequences questions.

Only the English script of the first scenario is shown in this chapter. All of the scripts in Dutch and English are shown in subsection B.2.

When the word <answer> is in the script, the robot will wait for an answer.

**Situation 1** In this situation, the child chooses between Peter, who is a good DJ but can get overstimulated in busy situations and does not give preference to being a DJ and Jan who really wants to be a DJ but is worse than Peter.

### Script

#### **During the game:**

The robot will instruct the child how to walk, how to use doors and will ask the child to read the learning goals. If the children fail to read the learning goal, the robot will read it for them.

#### **Question 1:**

The child chooses between someone who will be happy if chosen and someone who is considered the best person for the job.

Underlying values: Benevolence vs Achievement. Choosing Jan is considered as benevolence as he really wants to do the job. Choosing Peter is considered as the value Achievement as you will get the best possible situation.

Text before question: You made a good choice in choosing the DJ!

Clarification: Why did you choose him? <answer>

Implications & Consequences: How do you think Jan and Peter will feel because of your choice? <answer>

Text after question: Ah, I understand now why you chose him

#### **Question 2:**

The situation is the same as question 1 but in this case the hurdle of the best person not liking busy situations is removed.

Underlying values: Similar to question 1 but it takes away a hurdle.

Text before question: Would your choice be different if Peter would not have told you that he did not like busy situations? <answer>

Clarification: Why? <answer>

Implications & Consequences: What would Jan and Peter think about your decision? <answer>

Text after question: Thank you for your answer!

#### **Question 3:**

The child is now presented with the fact that choosing a DJ will result in the child that is not chosen, getting angry. The child must now choose if they are going to choose a DJ anyway or choose not to make a choice and let the teacher decide.

Underlying values: Power vs Security. Having the possibility to make a choice is considered as power. Knowing that you will make someone angry might risk your own security and stability of

your friendship.

Text before question: You know before you make the choice that the other person is going to get angry anyway. Would you still make a choice or would you go to the teacher and tell her to choose for herself <answer>

Clarification: Why did you choose this? <answer>

Implications & Consequences: What would happen if you do this? <answer>

Text after question: Yes, I understand why you choose this!

#### **Question 4:**

In this situation, the child is friends with one of the persons and the other person is considered best for the job. The dilemma is choosing between a friend and the best person for the job.

Underlying values: Benevolence vs Achievement. Choosing the friend is positive for Benevolence while choosing the best person is best for Achievement.

Text before question: Suppose you are friends with Jan but you know that Peter is a much better DJ and wants to be a DJ just as much as Jan, would you choose your friend or the one who is better? <answer>

Clarification: Why did you choose this? <answer>

Implications & Consequences: How would Jan and Peter respond to this choice? <answer>

Text after question: I understand you a little better now!

#### **End:**

Great, we are done with the first level, do you remember the two learning goals?

(comment: the participant either answered yes (they remembered), no/forgot (they did not remember), or directly gave the good answer without answering yes or no)

Yes: Wow, you are doing great, can you tell them to me?

Good answer: Wow, you do not even need me!

Wrong answer: It is going great! I remembered them as well, I will tell them to you.

The first is: Some children can handle busy environments a bit worse than others, they might get overstimulated

The second is: Some children have a bit more energy to expend than others

No/Forgot: Okay, no problem, we are learning together and I remembered them as well, I will repeat them: The first is: Some children can handle busy environments a bit worse than others, they might get overstimulated. The second is: Some children have a bit more energy to expend than others

Good answer: Wow, you do not even need me!

We are now done with the first couple of questions, you can click on done in the game.

## **5 Evaluation Method**

This chapter presents the evaluation method to test the design of the robot of subsection 4.1, i.e., the research questions of subsection 2.5. *H1a: The “Clarification” type of Socratic Questioning by the*

robot will lead to a better (value-based) argumentation quality of the child than the Socratic Questioning type of “Implications & Consequences”.

*H1b: A recurrence of a Socratic Question will lead to a better (value-based) argumentation quality than the previous Socratic Questions.*

*H2: There exists an interaction effect between the recurrence of the questions and the type of question that results in a significant impact on the quality of argumentation of the participant*

*H3: The personality traits of a participant have an impact on the quality of argumentation of the participant*

If we combine the three hypotheses, we theorize that there is a direct effect on the score based on the type of question and on the recurrence of the questions, there is an interaction effect between the recurrence and type of question, and there is a direct effect based on the personality of the participant. The suggested conceptual framework is shown in Figure 1

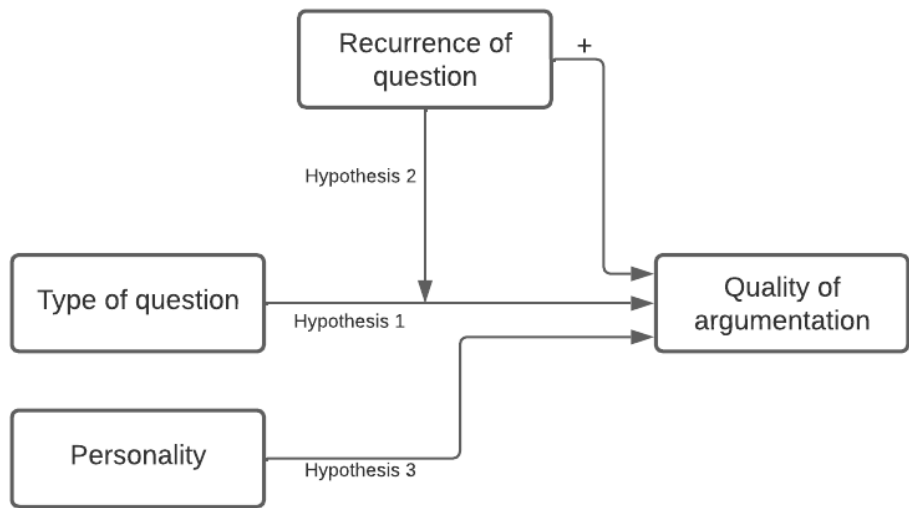


Figure 1: The conceptual framework of independent and dependent variables. The quality of argumentation is the dependent variable. Recurrence, type of question and personality are independent variables. Recurrence is also a moderator variable on the type of question. The recurrence is thought to have a positive impact on the quality of argumentation.

## 5.1 Participants

A school was contacted through the Groep "Grade" 6 teacher to participate in the study. All students in the classes were eligible to participate as they are all Dutch students and there were no visual or hearing impairments.

The test participants were children within the age range of 9 to 11. In total 26 students participated. For one participant, the audio-recording did not succeed. The data from this participant is removed. The school is situated in a neighbourhood with a higher proportion of higher-educated people.



## 5.2 Experimental design

The participants are divided equally into two groups. Both groups are getting the same levels and dilemmas. The difference between the two groups is that group 0 will get a Clarification question at the first dilemma and group 1 will get a Implications & Consequences question at dilemma one. At the next dilemma, the type of questions will be reversed. This continues for each of the sixteen dilemmas. The participants are getting sixteen questions, eight questions per Socratic Questioning type. This ensures that the experiment is within-subject, and the order of condition is counter-balanced.

As mentioned, the participants are getting sixteen dilemmas and accompanying questions over the time span of the experiment. As the questions are repeated, the experiment uses repeated measures.

The dependent variable is the quality of argumentation. This quality of argumentation can depend on the type of question, on the recurrence of the question and on the personality of the participant.

## 5.3 Assignments for participants

In the session, the participants make decisions based on moral dilemmas. These moral dilemmas are introduced in the game, shown in subsection B.1. The dilemmas are further elaborated on by the robot, shown in subsection B.2. The participants are also introduced to learning goals and asked to remember them.

## 5.4 Variables

Variable	Timing	Measured	Type
(1) Opinion about robots	Before session	Pre-session questionnaire	Interval
(2) Big 5 personality traits	Before session	Pre-session questionnaire	Ratio
(3) Number of learning goals remembered (during)	During session	Audio-recording	Ratio
(4) Graded answers to each question	During session	Audio-recording	Interval
(5) Number of learning goals remembered (after)	After session	Post-session questionnaire	Ratio
(6) Opinion about the experiment robot	After session	Post-session questionnaire	Interval
(7) Position of Schwartz' values	After session	Post-session questionnaire	Ordinal

Table 2: Overview of all variables. The variables are ordered based on occurrence in the experiment

### 5.4.1 Pre-session

An initial questionnaire has been done, this questionnaire is shown in subsection A.1. This questionnaire occurred before the session with the robot.

#### 5.4.1.1 (1) Opinion of robots

The participants were asked about their opinion of robots in general. They were asked if they thought robots were cool, scary and exciting. The questions that map the opinion about robots were done using five possible answers: Not at all, a little bit, neutral, very, very much.

#### 5.4.1.2 (2) Big 5 personality traits

Some questions were asked that aimed to map the participants' personalities. The method used was using the questions from the Pictorial Personality Traits Questionnaire for Children (PPTQ-C) [55]. However, the pictures were not used as it was not possible to import these into the survey software,

only the questions that accompanied the pictures were used. In total, seventeen questions about the personality of the participant were asked. These questions introduced a statement and asked the participant if this occurs never, sometimes, on average, often or very often.

The personality questions asked for the questionnaire related to the big 5 personality traits: extraversion, neuroticism, openness, conscientiousness, and agreeableness. Each of the questions corresponds to one of the traits. The answers to the personality questions were reshaped to integers from zero to four. The first option in the Likert scale which usually represented “Not at all” was configured to a zero. The last option was configured to a 4.

Some of the questions in the personality questions were asked in reverse, which means that the leftmost answer to the question did not correspond with a 0 but with a 4. The questions that correspond to a certain personality trait were combined to create a value that represents that trait. This is the method used in PPTQ-C[55]. The result for each participant were five personality traits with a value between zero and twelve.

1. Extraversion

- (a) (Reverse) I usually play alone
- (b) If others have fun, I join them
- (c) (Reverse) If someone is joking, I do not usually see why it is funny

2. Neuroticism

- (a) Usually, when I go to school, I am worried
- (b) If something goes wrong, I get nervous quickly
- (c) I am often worried about something

3. Openness

- (a) I often notice birds flying when playing outside
- (b) (Reverse) On vacation I like to relax instead of discovering something new
- (c) I like to learn new and different things

4. Conscientiousness

- (a) I like doing my homework
- (b) My room is organized
- (c) When I get money, I save it

5. Agreeableness

- (a) I notice it when someone in my class needs something
- (b) When I see that I can help someone, I will help them
- (c) When I have a new toy, I share it with others

#### 5.4.2 During session

The participant was audio-recorded during the entire session. This recording was transcribed.

Rating	Explanation	Context of example	Value related to example	Example
0	No answer, no argumentation	-	-	"I don't know"
0.25	Explanation without value, bad quality of argumentation	Why would you step in between the fight?	Power	"Because six year olds can also have fights"
0.5	Brief mention of value, moderate quality of argumentation	Why would you choose the better person instead of your friend?	Achievement	"Because you don't always have to choose your friend"
0.75	Hidden insight in value, medium quality of argumentation	Why did you let the teacher choose instead of choosing yourself?	Achievement	"I would tell the teacher because she can always make great decisions"
1	Clear insight in value and good quality of argumentation	Why did you choose Jan over Peter?	Benevolence	"Peter said he thinks it is difficult for him and Jan really wanted to be DJ"

Table 3: Rating system for argumentation, An example answer is shown with the context of the example and underlying value

#### 5.4.2.1 (3) Number of learning goals remembered (during)

During the session, after each level and the questions corresponding to that level were done, the robot asked if the participant remembered the learning goals. The participants could answer with yes, no or by directly stating the learning goals. The number of learning goals remembered is summed for each participant. If they answered yes, the robot would ask them to tell them the learning goals.

#### 5.4.2.2 (4) Graded answers to each question

The answers to each of the questions related to the dilemmas were rated a score from zero to one based on the quality of the argumentation. To analyze the quality of the argumentation, the answers are first given a score. An answer where there is a clear insight into values with a good quality of argumentation will result in the best possible score of 1. An answer that has some hidden insight into values will result in a score of 0.75. When the answer describes a value briefly it will result in a score of 0.5. Answers which have an explanation without value will give a score of 0.25 and no answer or an "I do not know" results in a 0. The rating system is summarized and examples are shown in Table 3. The grades are not exclusive so multiple answers could have the same score. A list is made per question consisting of all answers from the participants which allows for the answers of the participants to be compared to the answers of their peers. When necessary, the score of answers got adjusted if the score was deemed incorrect when compared to answers from their peers. This ensures that similar answers receive a similar rating.

The grades could be considered as being subjective, to reduce this, another researcher was asked to rate ten random answers from the participants. Some joint training was first performed on part of the data to learn the rating system. Then the other researcher was given different data to rate. The other researcher gave the exact same rating in six of these ten answers. In one answer, the rating was one step lower due to the assumption being "easy". The other answer was slightly higher as the other researcher showed that the participant had a clear value for authority, the connected value was Achievement which is why the given rating was slightly lower. The other researcher gave similar grades with a slightly different explanation, showing that rating the answers still is very subjective. To further check the accuracy of the ratings, the Fleiss' Multirater Kappa [56] has been calculated. In Figure 2, it is shown that the Fleiss' kappa is .433. So the agreement is above chance agreement. Landis and Koch consider this a moderate strength of agreement. We can also consider the results as statistically significant as the p-value is .017 which is below .05.

Overall Agreement <sup>a</sup>						
	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
			z	Sig.	Lower Bound	Upper Bound
Overall Agreement	,433	,181	2,392	,017	,078	,787

a. Sample data contains 10 effective subjects and 2 raters.

Figure 2: Table showing the Fleiss' Multirater Kappa

The average of all the answers provided by each participant in response to the robot’s questions was computed. In addition, the average of all of the Clarification and Implications & Consequences answers were computed per participant. The averages per level were also computed and for the second level, the average was also calculated without the answers to question 6 (level 2 question 2) as explained in subsection 6.1.

### **5.4.3 Post-session**

After the experiment, the participants would get a second questionnaire. This questionnaire can be found in subsection A.2.

#### **5.4.3.1 (5) Number of learning goals remembered (after)**

The participants were asked in the questionnaire what they learned from each level; which learning goals they remembered. The number of learning goals remembered was summed.

#### **5.4.3.2 (6) Opinion about the experiment robot**

In the questionnaire, the participants were asked about their opinion of the experiment robot. The questions about the experiment robot were related to if they thought the robot was smart, fake, nice and like a human. The opinion about the experiment robot was again done using five possible answers: Not at all, a little bit, neutral, very, very much. Similar to the Pre-session questionnaire data, the answers to the opinion about the experiment robot were reshaped to integers.

#### **5.4.3.3 (7) Position of Schwartz’ values**

Finally, the participant was asked to assess the relative importance of the values mentioned in subsection 2.2 to them. The participant was shown two randomly selected values from the list of Schwartz’s ten values and was asked which one is more important to them. The chosen important value was put above the other value and an additional value was introduced. The participant was asked to place this value on the table to represent their perceived level of importance, which could be above the top value, in the middle or below the lowest value. This process was repeated for each value until all ten values were ranked. For the data on the participants’ values, there was a score given based on the position that the participant gave the value. If the participant placed the value on position one, so the most important value, the score would be 1. If the participant placed the value on position ten, the least important position, the score would be 0,1.

## **5.5 Material and Experimental setup**

The robot used was an NAO V6 with Interactive Robotics Software installed. The robot was being controlled by a Lenovo Yoga. The robot was connected to the “Robots in de Klas” server. The robot was controlled in a Wizard of Oz manner, as explain in subsection 4.4. Due to the robot already being controlled in a Wizard of Oz manner, there was a lock where the robot would wait for the prompt from the Wizard of Oz page. Due to this, it was not possible to connect the game to the robot. Certain occurrences in the game that would prompt conversation from the robot would be initiated from the Wizard of Oz software by the researcher.

The participants were playing the game on an HP Omen. The game has been made with Unity and the audio recording was done with Audacity.

In the room, the researcher was present. The participant had their own desk with a laptop and mouse and there was a NAO robot on this desk. The researcher and a research assistant were sitting behind the participant. The researcher was controlling the robot on a separate laptop and the research assistant was recording any irregularities in the session. There was a microphone in the laptop facing the participant that recorded the audio in the room. This audio has been transcribed and anonymized and the audio file has been deleted. The setup with a participant is shown in Figure 3.



Figure 3: Experimental setup. The participant has just finished the first level and is reflecting on the level and their values with the robot. The researcher is sitting behind the participant. The researcher's laptop is visible at the bottom of the figure.

## 5.6 Procedure

Information and consent forms were created for both the children and parents/guardians. The consent forms for the children are adapted to be understandable for children. The consent forms can be found in Appendix C. The parents were given the consent forms a week before the experiment. A week in advance, a schedule has been made for two weeks where all students can participate in a one-on-one session with the robot, assumed was that all children would participate. Some leeway had been created in the schedule if this would not be the case. Fourty minutes were planned per participant to calculate for movement to and from the classroom and reset time. The participants were separated into two separate groups. Which condition participants got was random.

The children were given the consent form when they arrived at class on the first day of the experiment if their parents had handed in the consent form. When all the applicable children had handed in the form, the robot was introduced in the class to remove any potential nervousness of the children about the robot.

The children who were able to participate were given the pre-session questionnaire. The participant filled in the questionnaire while all sitting in the same classroom. The questionnaire took about ten minutes and was filled in on the Chromebooks that were provided by the school. The variables gathered in this questionnaire were: **(1) Opinion about robots** and **(2) Big 5 personality traits**.

After this the sessions started. Each session took approximately half an hour. This consisted of playing the game for twenty minutes and doing the questionnaire for another ten minutes.

The participant was asked to play the game on the laptop. The design of this game is explored in section 3. The game includes learning goals, which are explored in subsection 3.3. The audio recording started when the participant walked into the room. The participants were first asked to create themselves in the character builder. When they mentioned that they finished building the character, the robot would turn on. The robot would introduce themselves, ask about the name of the participant and about the favourite food of the participant. The robot would make a joke to ease the tension. Next, it explained to the participant what would happen in the session. The robot would now tell the participant to start the game.

When a new learning goal would appear on the screen, the robot would ask the participant to read it to them. If they were unable to read the learning goal, the robot would read it to them.

In the game, participants were confronted with a certain scenario, and during each scenario, a decision was prompted for the participant. When a certain decision had been made, the robot would talk with the participant about the decision. Then the robot would give a bit of context, introducing a new but similar dilemma and once the decision at this dilemma had been made, the robot would inquire about their decision. In total, three additional decisions per game level were asked. After these questions were asked, the robot would ask if the child retained the learning goal. If they forgot the learning goal, the robot would assist by repeating the learning goal. The variable connected to these learning goals is: **(3) Number of learning goals remembered (during)**.

The reflection of decisions were either questions that ask for Clarification or questions that probe Implications & Consequences. Which type of question the robot asked, would depend on the condition group that the child had been placed in. In group 0, the robot asked a Clarification question on the first decision. In group 1, the robot asked the child a question that probed Implications & Consequences. Then on the next question, it was switched around. This allowed the test to be done within-subject. The answers of the participant to these questions relate to **(4) Graded answers to each question**. The child played four levels with in total sixteen questions.

After the participant finished the game. They were asked to fill in a questionnaire. This questionnaire first tested the regained learning goals. This relates to **(5) Number of learning goals remembered (after)**. Finally, there were some questions regarding the opinion of the children about the experiment robot. This related to: **(6) Opinion about the experiment robot**. Finally, there were also questions about the perceived importance of values for the participant. The participants were given two randomly selected values out of Schwarz's ten values and are asked to place them in a list based on their perceived importance. The next randomly selected value is given to the participant to place on their list. This is continued until ten values are rated on importance. After this, the audio recording was stopped and the participant was thanked and the participant returned to the classroom.

The audio recording is transcribed and the data is processed according to the method explained in subsection 5.4. After the pre-processing, statistical tests were run.

## 5.7 Statistical tests

To test if the data is normally distributed, a test for normality has been to be performed. The performed test is the Kolmogorov-Smirnov and Shapiro-Wilk test of Normality. This test is an appropriate test for sample sizes under fifty. As the sample size for this test is twenty-five. To test if outliers exist in the data, the Tukey's Hinges test is performed.

Hypothesis one a is comparing the averages of the Clarification and Implications & Consequences. To compare these two averages, a one-Sample t-test is performed. Next, the questions are individually compared with a Compare Means Independent Samples t-test.

Hypothesis one b checks if the recurrence of a Socratic Question leads to a better (value-based) argumentation quality than the previous Socratic Questions. To compare this, a Univariate Analysis of Variance is run to see if the recurrence of questions leads to a significantly better quality of argumentation.

The second hypothesis analyses if there is an interaction effect between the type of question and the recurrence of the questions. To test this hypothesis a Univariate Analysis of Variance with Repeated Measures is used.

Finally, the third hypothesis says that the five personality traits have an influence on the quality of argumentation. To analyse this, a stepwise regression is done. A regression test is done with all five traits, the least significant trait is removed, the test is run again and this is repeated until the most significant personality trait is left.

## 6 Evaluation Results

In this section, the results of the experiment will be shown, they will be analyzed per hypothesis. First some outliers will be discussed which are of impact on the data

### 6.1 Failed data collections

**Audio recording missing.** In the session with participant 8, the audio recording did not start. This resulted in the loss of all data from the audio recording. In the statistical analysis, all data from participant 8 was removed.

**Question six.** Due to the formulation of the question, many participants provided only yes and no responses to the implication and consequences question. The question proved to be a closed question instead of an open question. The question was: "Would the consequences of your choice be different



now?” Due to only getting yes and no answers, no values were represented and the resulting score was very low. The data from this question is considered unusable. Consequently, this question would not be considered significant for the final results. The data from this question is removed.

## 6.2 Hypothesis 1

The hypothesis which initially needs to be tested is the following: *H1a: The “Clarification” type of Socratic Questioning by the robot will lead to a better (value-based) argumentation quality of the child than the Socratic Questioning type of “Implications & Consequences”.*

Figure 4 shows the mean quality of argumentation scores for each question and questioning type. In this figure, one can see that the Clarification questions would often result in a slightly better **grade per question (5)** except for questions one and five.

Overall, the quality of argumentation score proved to be higher for the Clarification questions (Mean=.75) than the Implications & Consequences questions (Mean=.61), see Figure 5. The Kolmogorov-Smirnov and Shapiro-Wilk test of Normality showed that there is a normal distribution (for Clarification:  $p=.873$  and for Implication:  $p=.448$ ), and no outliers were detected. A t-test shows that the difference in argumentation quality for the two conditions is significant Figure 6.

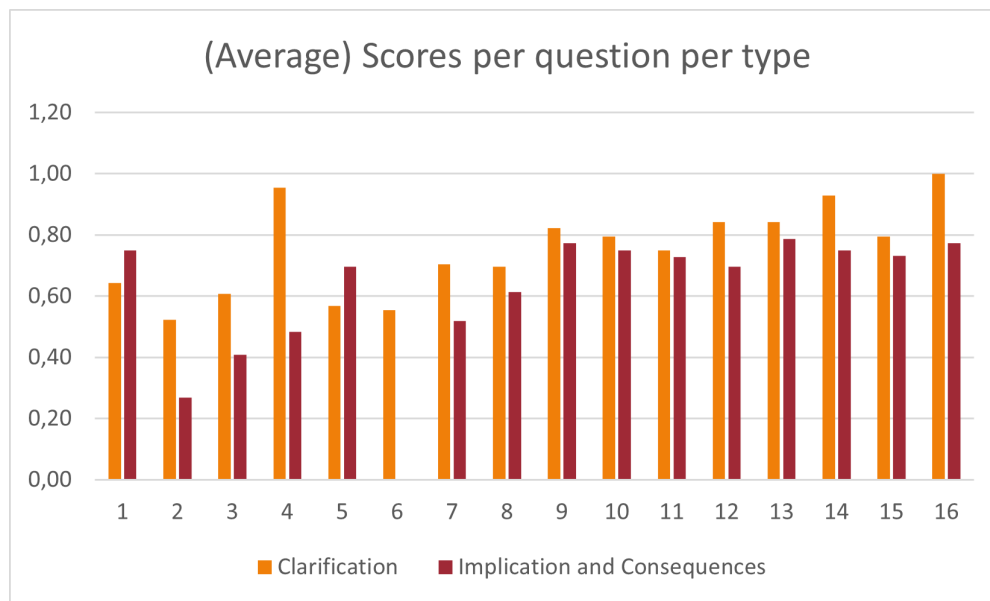


Figure 4: Measured is the average score per question per type of question. The horizontal axis represents all sixteen questions. The vertical axis represents the average score per question. The maximum score is 1,00.

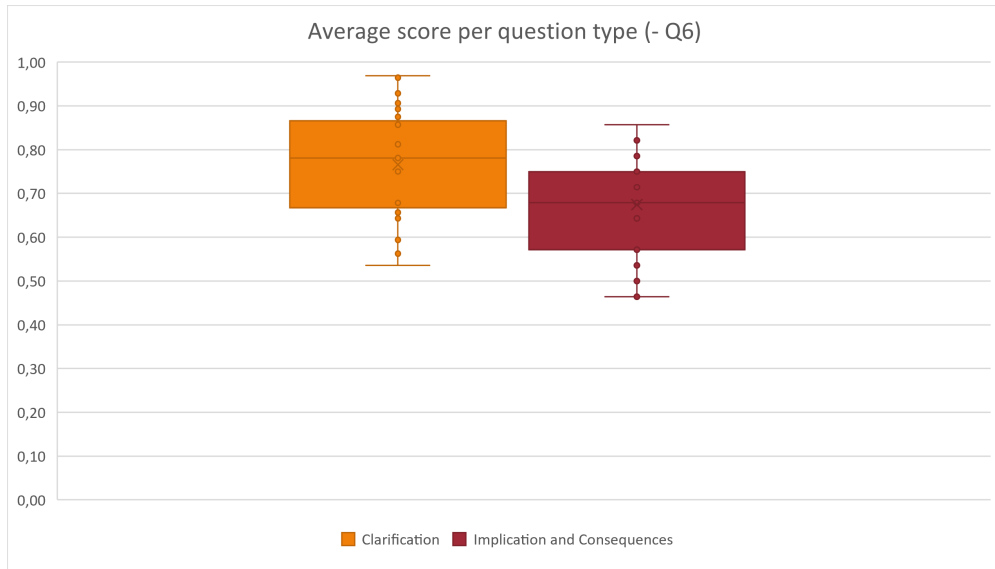


Figure 5: Measured is the average score per type of question. The vertical axis is the score. The maximum score per question is 1,00. Question six is removed due to the closed formulation of the question as mentioned in subsection 6.1.

**One-Sample Test**  
Test Value = 0

	t	df	Significance		Mean Difference	95% Confidence Interval of the Difference	
			One-Sided p	Two-Sided p		Lower	Upper
AverageClarification	40,608	24	<.001	<.001	,76680	,7278	,8058
AverageImplication	26,890	24	<.001	<.001	,64680	,5972	,6964

Figure 6: Table with t-test between Average Clarification and Average Implication and Consequences

The result of the t-test is shown in Figure 6. For AverageClarification the result is statistically significant ( $t = 40,608$ ,  $p < .001$ ). The results for AverageImplication are also significant ( $t = 26,890$ ,  $p < .001$ ). This allows us to accept the hypothesis; The type of question has an effect on the quality of argumentation. As there is a significant effect based on the type, based on Figure 5, we can say that Clarification results in a better quality of argumentation.

Now we will also see which of the sixteen questions will show a significantly different result for the Clarification or Implication and Consequences type of answers. This is done by Compare Means Independent Samples t-test. The test variable is the score of all participants at the fifteen questions and the grouping variable is the condition. If there is a significant difference, it means that a certain type of question is significantly better than the other.

Independent Samples Test											
		Levene's Test for Equality of Variances		t-Test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
Question 1	Equal variances assumed	3,040	,095	1,299	23	,103	,207	,10714	,08250	-,06352	,27781
	Equal variances not assumed			1,314	22,443	,101	,202	,10714	,08153	-,06174	,27603
Question 2	Equal variances assumed	2,175	,154	2,151	23	,021	,042	,25487	,11851	,00971	,50003
	Equal variances not assumed			2,038	16,081	,029	,058	,25487	,12503	-,01008	,51982
Question 3	Equal variances assumed	1,223	,280	-1,422	23	,084	,168	-,19805	,13928	-,48617	,09007
	Equal variances not assumed			-1,444	22,614	,081	,163	-,19805	,13719	-,48211	,08601
Question 4	Equal variances assumed	4,161	,053	5,220	23	<,001	<,001	,47240	,09050	,28520	,65961
	Equal variances not assumed			5,570	21,129	<,001	<,001	,47240	,08481	,29610	,64871
Question 5	Equal variances assumed	3,341	,081	-1,233	23	,115	,230	-,12825	,10404	-,34348	,08698
	Equal variances not assumed			-1,168	16,017	,130	,260	-,12825	,10984	-,36107	,10458
Question 7	Equal variances assumed	,153	,699	1,362	23	,093	,186	,18669	,13703	-,09678	,47015
	Equal variances not assumed			1,322	18,660	,101	,202	,18669	,14117	-,10916	,48253
Question 8	Equal variances assumed	,229	,636	-,574	23	,286	,571	-,08279	,14412	-,38093	,21534
	Equal variances not assumed			-,562	19,448	,290	,581	-,08279	,14740	-,39083	,22524
Question 9	Equal variances assumed	,069	,795	-,550	23	,294	,588	-,04870	,08853	-,23184	,13443
	Equal variances not assumed			-,557	22,443	,292	,583	-,04870	,08749	-,22993	,13253
Question 10	Equal variances assumed	,328	,572	,398	23	,347	,695	,04545	,11434	-,19108	,28198
	Equal variances not assumed			,388	19,385	,351	,702	,04545	,11701	-,19913	,29004
Question 11	Equal variances assumed	,000	,994	-,195	23	,424	,847	-,02273	,11677	-,26428	,21883
	Equal variances not assumed			-,192	20,534	,425	,849	-,02273	,11818	-,26883	,22337
Question 12	Equal variances assumed	2,107	,160	1,437	23	,082	,164	,14448	,10051	-,06345	,35241
	Equal variances not assumed			1,495	22,876	,074	,149	,14448	,09662	-,05546	,34442
Question 13	Equal variances assumed	,066	,799	,557	23	,291	,583	,05519	,09909	-,14980	,26019
	Equal variances not assumed			,564	22,514	,289	,578	,05519	,09780	-,14736	,25775
Question 14	Equal variances assumed	,036	,851	-1,958	23	,031	,062	-,17857	,09121	-,36725	,01010
	Equal variances not assumed			-1,864	16,579	,040	,080	-,17857	,09578	-,38103	,02389
Question 15	Equal variances assumed	,877	,359	,608	23	,275	,549	,06331	,10410	-,15203	,27866
	Equal variances not assumed			,615	22,392	,272	,545	,06331	,10296	-,15001	,27663
Question 16	Equal variances assumed	86,871	<,001	-2,631	23	,007	,015	-,22727	,08637	-,40594	-,04861
	Equal variances not assumed			-2,319	10,000	,021	,043	-,22727	,09802	-,44567	-,00888

Figure 7: Table with the t-test for every question per condition.

In Figure 7, there are four questions where a  $p \leq .05$  is reported: Questions 2, 4, 14 and 16. The questions where there is a significant difference are questions two, four, fourteen and sixteen. It does not seem to be the case that the Implication and Consequences scores of questions four, fourteen and sixteen are significantly lower than the surrounding Implication and Consequences questions. There is a significantly lower score for the Implications & Consequences question at question two.

Next, we will evaluate Hypothesis 1b: *H1b: A recurrence of a Socratic Question will lead to a better (value-based) argumentation quality than the previous Socratic Questions..* Figure 9 and Figure 10 show a significant effect ( $F(14,345)=5.62, p \leq .001$ ) of recurrence on the quality of argumentation.

### 6.3 Hypothesis 2

The second hypothesis is: *H2: There exists an interaction effect between the recurrence of the questions and the type of question that results in a significant impact on the quality of argumentation of the participant*

In Figure 8, the **grade that each participant got per question (5)** is shown. The grades are shown in a stacked line chart for clarity. There is a clear increase except for questions two and six. These questions are outliers as explained in subsection 6.1.

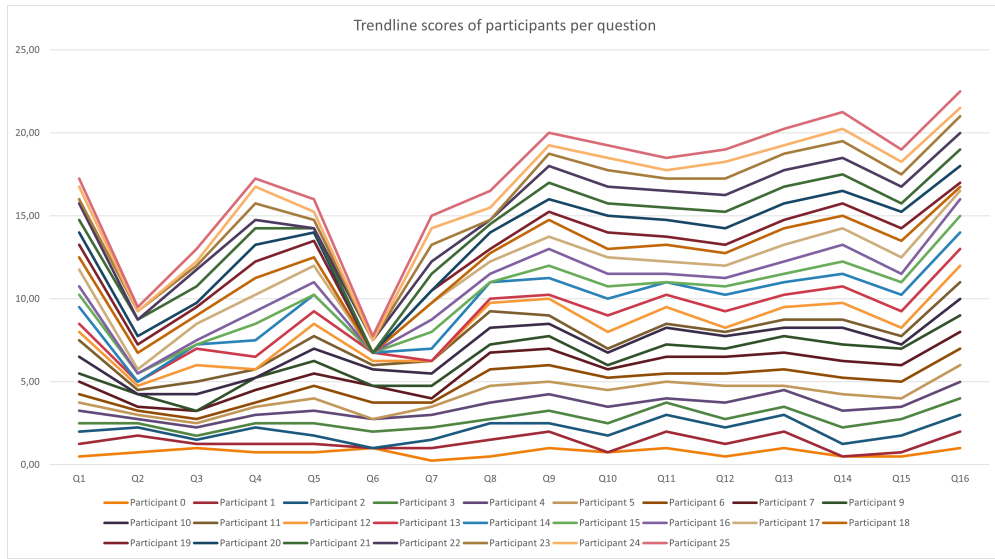


Figure 8: Measured the score per question. The horizontal axis represents all of the sixteen questions that the participants were asked. The vertical axis is the total score. Each of the twenty-five participants could score a maximum of 1,00 points per question thus the maximum per question is twenty-five.

As there is an increase in score visible, the recurrence of the question might be relevant. The recurrence might also have an impact on the type of question. To test if there is an impact of the recurrence, some statistical tests are necessary.

**Tests of Between-Subjects Effects**

Dependent Variable: Score

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	9,367 <sup>a</sup>	29	,323	4,372	<,001	,269
Intercept	184,451	1	184,451	2496,750	<,001	,879
Type	1,251	1	1,251	16,929	<,001	,047
Recurrence	5,811	14	,415	5,619	<,001	,186
Type * Recurrence	1,912	14	,137	1,849	,031	,070
Error	25,487	345	,074			
Total	221,063	375				
Corrected Total	34,854	374				

a. R Squared = ,269 (Adjusted R Squared = ,207)

Figure 9: Table with the Univariate Analysis of Variance with Repeated Measures results

The results of this test can be found in Figure 9. The results of the Univariate Analysis of Variance showed a significant effect of the type of question on the score per question ( $F(1,345)=16.93$ ,  $p \leq .001$ ), indicating that the type of question was associated with a difference in score. There was also a significant effect of the recurrence on the scores per question ( $F(14,345)=5.62$ ,  $p \leq .001$ ), indicating that recurrence was associated with an improvement in score.

However, these effects were moderated by an interaction between the type of question and recurrence ( $F(14,345)=1.85, p=.031$ ). This means that the effect of the type of question on the score varied depending on the recurrence.

The profile plot for this interaction is shown in Figure 10. The lines cross each other three times in the first half of the image, indicating an interaction effect mainly in the first half of the experiment.

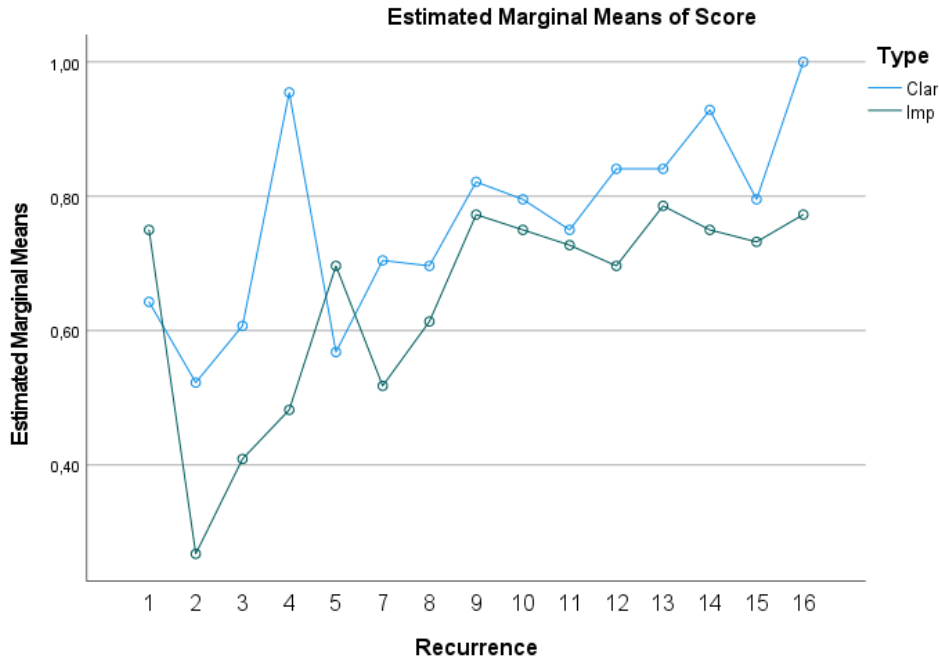


Figure 10: Graph with the profile plots for Clarification and Implication and Consequences

As an interaction effect exists, we can accept the hypothesis. There is an interaction effect between the recurrence and the type of question that significantly impacts the quality of argumentation.

### 6.4 Hypothesis 3

The third hypothesis is the following: *H3: The personality traits of a participant have an impact on the quality of argumentation of the participant.* To test if the hypothesis is accepted, the data of the personality is visualized first, then a statistical analysis is done.

In Figure 11, the personality traits of the participants are visualized. Each participant is represented on the circular axis. The personality traits are a line. In this figure, it is clear that extraversion is most often on the outside axis and thus is very prevalent in the participant group. Neuroticism is by most participants the least prevalent trait. In Figure 12 the same data is flipped so the personality traits are on the circular axis. This allows for an average representation of the participant group, this is the blue shape. As was also clear in the previous figure, extraversion is the highest, neuroticism is the lowest and the other three traits are average. Figure 13 shows a boxplot with every personality trait. This allows for the insight in the quartile ranges and the means of the personality traits.

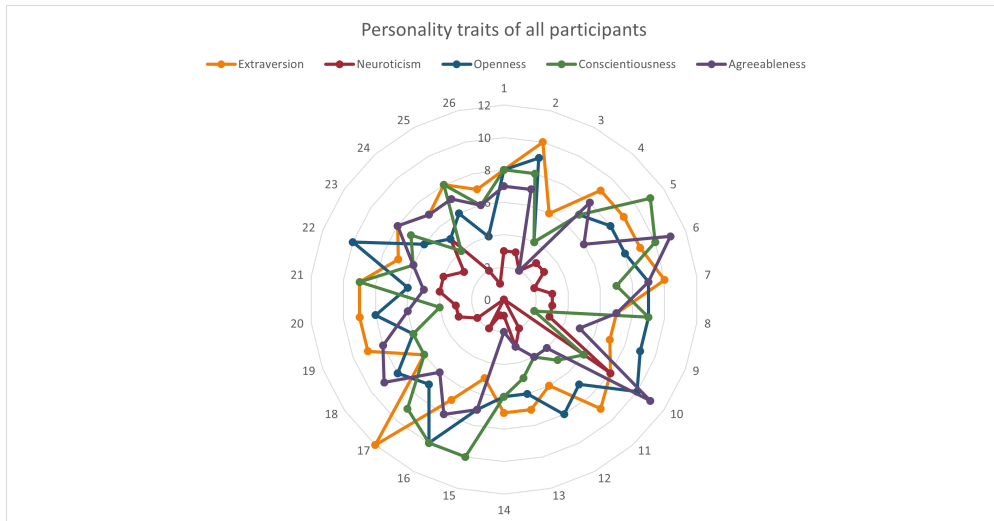


Figure 11: Measured the Big Five personality traits in the participant group. The circular axis represents the twenty-six participants. The lines represent the five personality traits. The vertical axis is the intensity of the personality trait. There are four points to be gotten from each question which ensures a max amount of points from twelve.

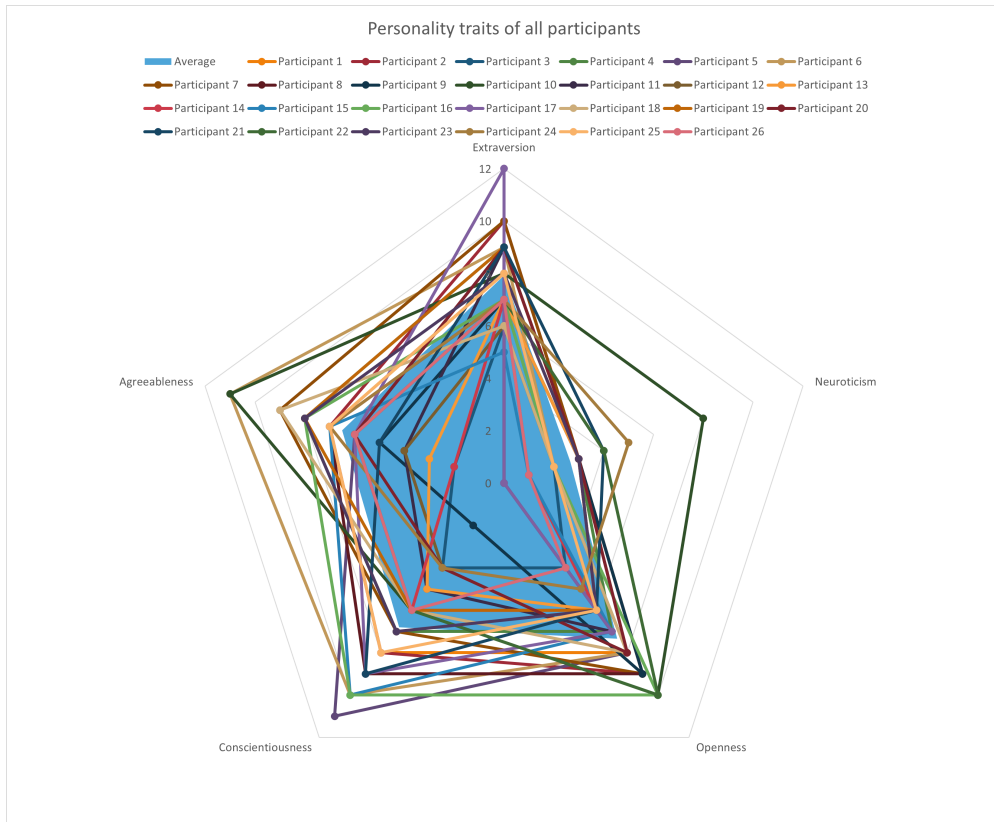


Figure 12: Measured the Big Five personality traits in the participant group. The circular axis represents the five personality traits. The lines represent the twenty-six participants. The vertical axis is the intensity of the personality traits. There are four points to be gotten from each question which ensures a max amount of points from twelve.

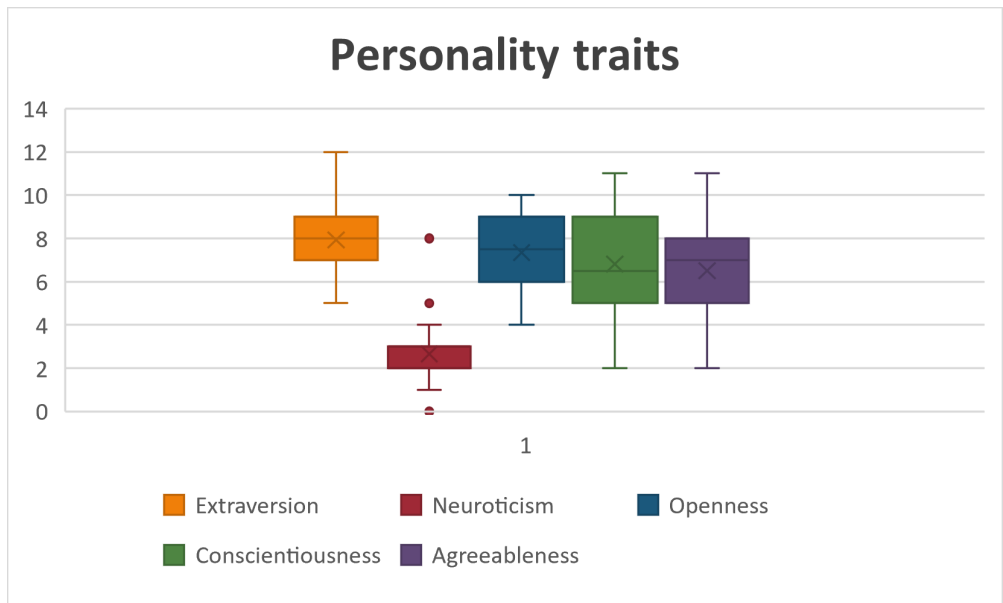


Figure 13: Measured the Big Five personality traits in the participant group. The horizontal axes show the different personality traits. The vertical axis shows the strength of the personality trait.

From the visualizations, it becomes clear that there is variation between the personality traits of the participants which means that a statistical analysis can be done to see if any of the personality traits have an impact on the quality of argumentation.

To see if any of the personality traits have an effect, stepwise regression is done. To perform this stepwise regression, regression is first run with all personality traits as independent variables. This is shown in Figure 14. Then the variable that explains the least significance is removed. This is Agreeableness. The regression test is run again with the remaining four traits, the least significant one is removed and this is continued until the most significant trait is remaining. This is shown in Figure 15.

Coefficients <sup>a</sup>						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,842	,105		7,986	<,001
	Extraversion	-,016	,009	-,383	-1,745	,097
	Neuroticism	,009	,012	,192	,772	,450
	Openness	-,009	,012	-,193	-,770	,451
	Conscientiousness	,010	,009	,265	1,063	,301
	Agreeableness	,000	,009	,012	,046	,964

a. Dependent Variable: avgmin6

Figure 14: Table showing the effect of the personality traits on the average score



**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,853	,083		10,230	<,001
	Extraversion	-,015	,008	-,355	-1,821	,082

a. Dependent Variable: avgmin6

Figure 15: Table showing the effect of Extraversion on the average score

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,355 <sup>a</sup>	,126	,088	,07580

a. Predictors: (Constant), Extraversion

Figure 16: Table showing the multiple correlation coefficient and coefficient of determination of Extraversion.

The stepwise multiple regression predicted the quality of argumentation based on Agreeableness, Extraversion, Neuroticism, Conscientiousness and Openness. Extraversion explained the most variance but did not have a significant impact.

As is clear in Figure 15 and Figure 16, The personality does not have an effect on the average score ( $t=-1.821$ ,  $p=.082$ ,  $R^2=.355$ ). As this is the most significant trait, we reject the hypothesis.

## 6.5 Qualitative Analysis

During the experiment, some observations have been made. These observations provide further insight into the child's behaviours and argumentation.

**Explanation at the choice question.** During the session, some participants gave an explanation of their decision without being prompted by the robot. For example, when they were asked if they would choose to step in between a fight or get the teacher, the answer would be: "I would step in between to ensure that they will stop fighting directly". This would interfere as some participants said that they already gave their answer to the question. Luckily, this happened only two times. Most often, the participants would slightly alter their answers to answer the question that the robot asked. The participants seemed to know that the robot would ask for the reason why they chose the option and they answered it before the robot could ask. This increased slightly over the course of the experiment. From the audio-recording one can extract when the participant gave an explanation before the question. The number of participants who gave an explanation without being prompted is shown in Figure 17.

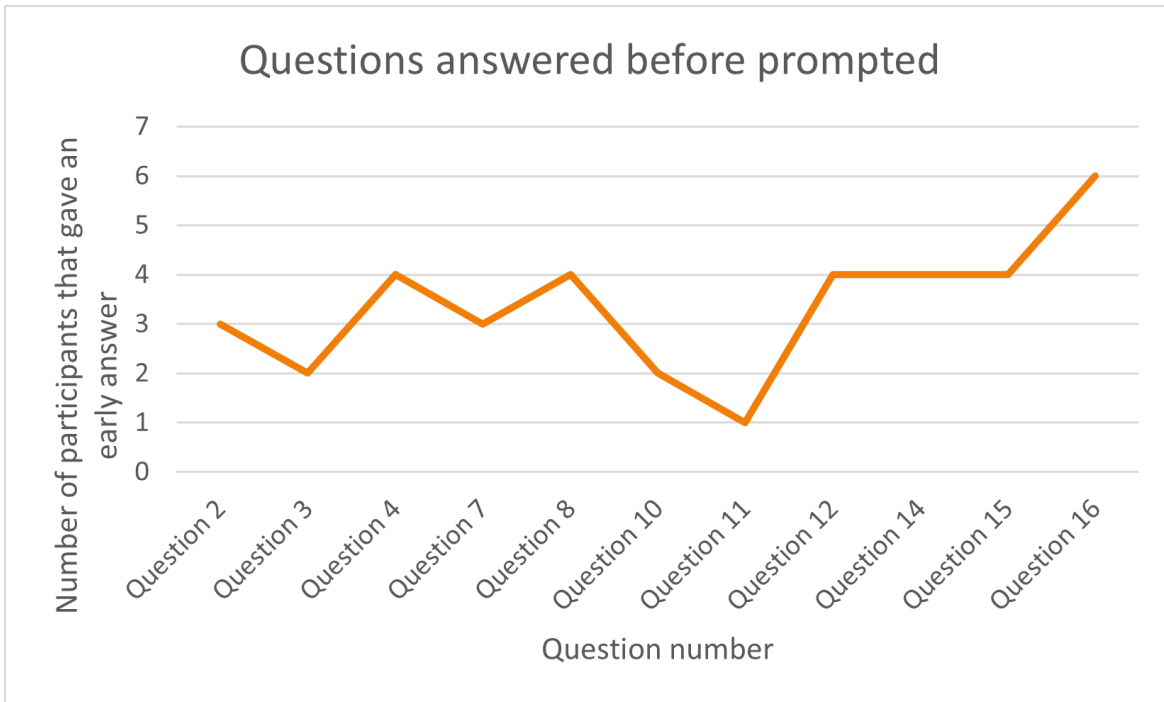


Figure 17: Amount of participants who gave an explanation without being prompted per question

Some translated quotes from the participants are shown. These quotes illustrate some examples where the participants gave an explanation before being prompted.

In the dilemma where the participant chooses if they will tell the teacher that the other student (Sasha) asked them to ask a grammar question to the teacher, the robot asks the participant: “Would you tell the teacher that Sasha asks the question?”. Participant 0 said: “If you would say her name, and Sasha is afraid to ask it, it will not be very nice.”. Participant 0 gave an explanation before the Socratic Question about their reasoning.

At the question of whether the participant would choose someone better for the job (Peter) or the one that wants to do the job (Jan) but where the hurdle of getting overstimulated is removed. The robot would ask: “If Peter would not get overstimulated and would be better for the job would you choose the better person for the job or the one who wants to do it the most?”. Participant 17 said: “Yes, I would choose Peter I think, because he is also really good.”

#### Swap to another value.

The participants were asked to choose between two options for every question. Some answers that the participants gave did not align with the values that were connected to these options but they would give an answer aligning with a different value.

For example, they were asked by the robot: Do you want to do something fun or something you would learn from? The two values connected to this question were hedonism and stimulation. Participant 0 answered: “I would let Mike (the teammate) choose because I would like both options”. The participant either showed some regard for Mike or would choose to not choose an option. Their answer was closest to benevolence and thus represented a different value than initially connected to this question.

Another example was whether the participant would ask the teacher to choose to prevent the persons

from getting angry at them. Participant 11 answered that they would choose the teacher which aligned with Safety. Their answer however was: “I would choose the teacher because the teacher can always make good decisions”. The explanation that they give is more aligned with Achievement as they prefer to make the best decision.

**Aligning with self-reported values.** The participants were asked to rank the list of **Schwartz’ ten values (8)** from most to least important. The answers given in the dilemmas did not always align with these self-reported values.

In Figure 18, the results of the self-ranking of the values are shown. The values are distributed from left to right based on the total score. As mentioned in ??, the answers are rated based on the position given by the participants. If these answers are combined, a total score per value appears. The values are shown from highest to lowest. Next, the position of importance that each value is given is shown. For example, benevolence has eleven first-place positions, eight second-place positions, two third-place positions and four fourth-place positions.

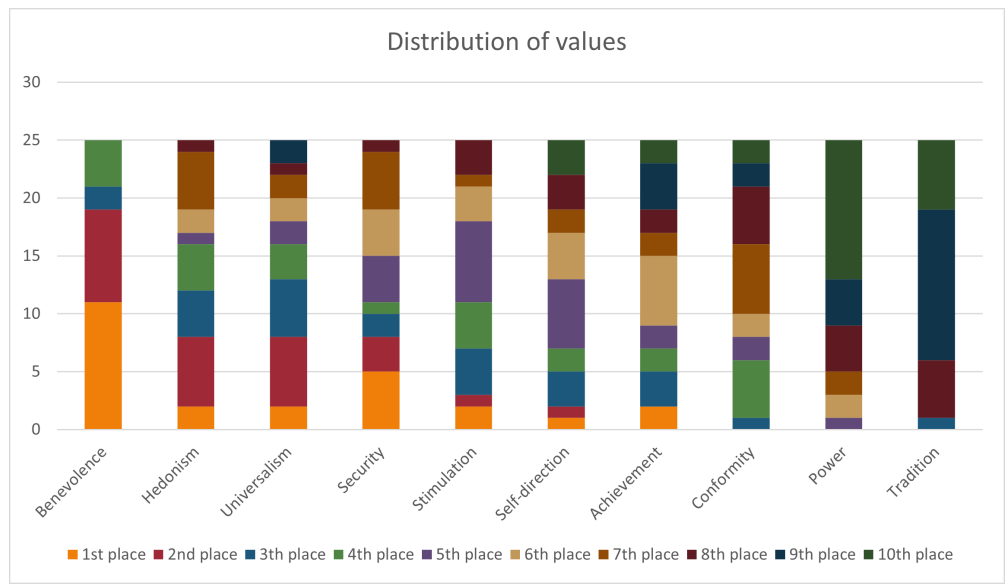


Figure 18: Measured the order of importance for the participants of Schwartz’ values. On the horizontal axis are the values. They are ordered from most to least important for the group of participants.

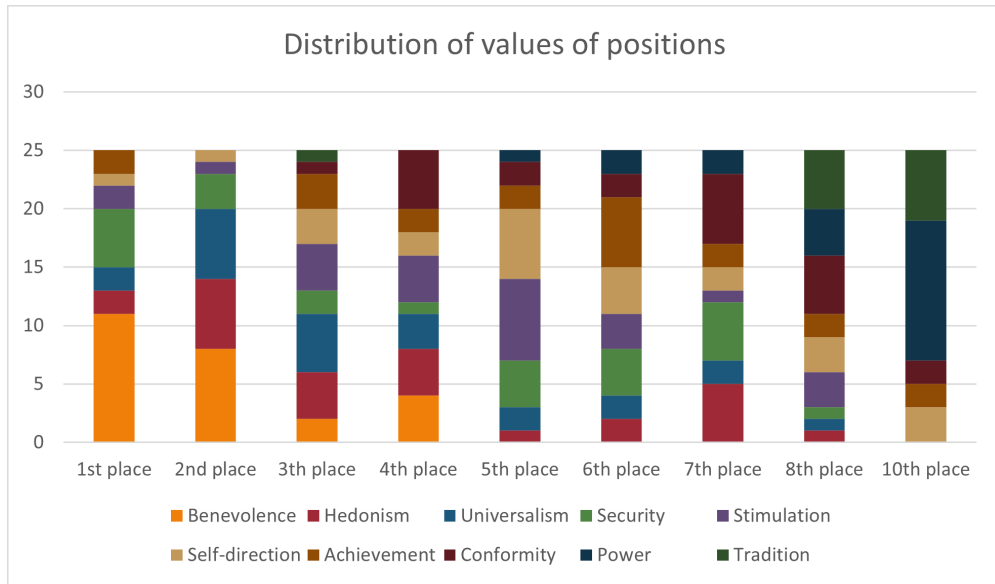


Figure 19: Measured the order of importance for the participants of Schwartz' values. On positions the values could get.

The answers that the participants gave to the questions did not always align with their self-reported importance of values. The order that the participants gave in the questionnaire in ??, often had benevolence in a very high position. During the experiment, participants often chose another value above the benevolent choice. This observation prompted a new method of looking at the gathered data: the data from the experiment could be used to show which questions had the highest rate of participants aligning with their self-reported values, as shown in Figure 20. And the percentage of questions that the participants answered aligning with their values, in Figure 21.

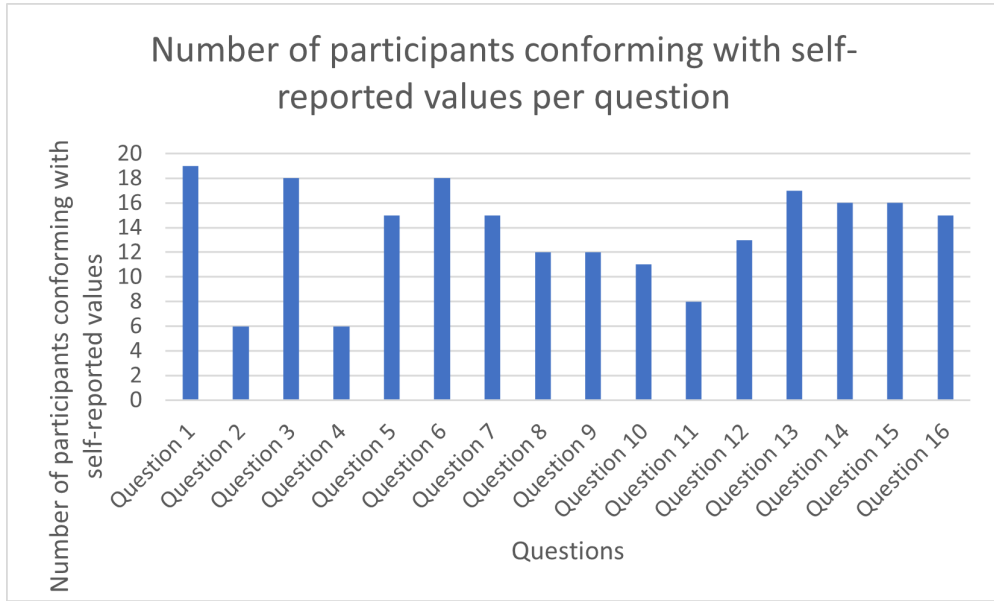


Figure 20: The horizontal axis demonstrates the questions while the vertical axis shows the number of participants whose answers to the question aligned with the self-reported order of importance.

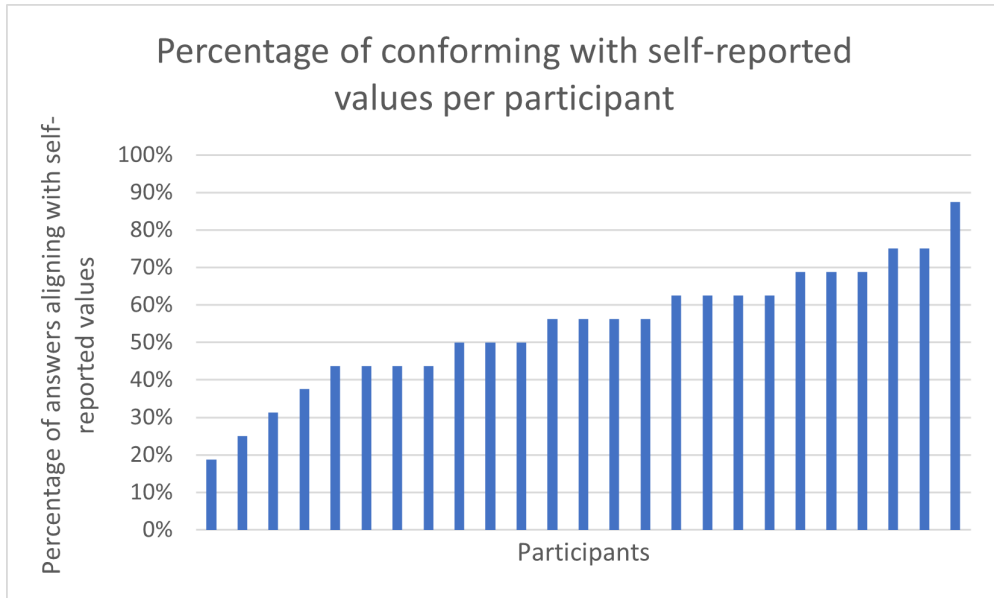


Figure 21: The horizontal axis demonstrates all participants ordered on the score. The vertical axis shows the percentage of answers that aligned with the self-reported order of importance.

Figure 20 shows that questions two and four have a very low rate of conforming with the self-reported order of importance of the values. These two questions compare achievement and benevolence and the participants often chose achievement over benevolence. This is interesting because benevolence was rated as very important, often taking the first or second position in the list of importance.

Figure 21 illustrates the score of conforming with the self-reported values. On average 54% of the answers aligned with the order of importance.

There was even one participant who answered a certain question with a certain value and would switch directly in the following question. The first question this participant got asked, is the one where the participant would know that Jan or Peter would get angry with the participant if the participant were to choose, and the participant is asked whether they would choose or let the teacher choose. Participant 13 said: “If they want to both (be the DJ) then I think it’s unfair and we should rather do something that is fairer because when someone is my friend, I would always choose them.” The next question was: Would you rather choose your friend or the better person for the job? The participant answered: “The better person for the job”. The first answer did align with their self-reported value but the second answer did not.

As forty-six per cent of the answers did not align with the self-reported order of importance of values, we can conclude that the decisions that are made when context is provided, are often not aligned with the order of importance of values that we believe are important to us.

There are two reasons that could explain this. First, value relevance is dependent on the context. Each question represents a different context. Second, putting benevolence high on the list might be a “desirable answer” as it is viewed as a desirable characteristic by society. Participants might have been raised with the notion that being kind to others is the most important but might not always agree with that subconsciously.

**Noticing the types of neurodiversity.** As the learning goals were very generic to prevent enhancing stereotypes, the types of neurodiversity were not explained to the participants. However, one participant noticed that the learning goal about the fact that some people had more energy to expend than others, was about ADHD. The participant explained that they have ADHD and thus recognized that the learning goal is about them.

**Willingness to answer questions.** All participants were taken out of class to perform the experiment. A similar experiment was done to find the values of children during a robot vacation camp. It was noticed that most of the children during the vacation camp were very eager to return to the camp and were less willing to talk. This resulted in some complaining that they were bored. The fact that the children in this experiment were taken out of class, made them more willing to talk.

In the robot camp, the children often mentioned that there were too many questions and it felt repetitive. This often occurred around the fifth or sixth question. In this experiment, the participants answered sixteen questions in total. None of the participants complained about the number of questions. This could have occurred due to either the situation of being in school instead of in a vacation camp or due to the game in between the sets of questions. There have been some reports that gaming influences motivation [57] which is why it is a reasonable suggestion to say that the game improved the motivation. In addition, due to the game, the participants had some variation between only answering similar questions. It is suspected that due to the game, the participants were less likely to get bored. However, more research is needed to find the true influence of a game in between sets of questions.

**Opinion of robots and NAO** The participants had a very positive opinion about robots and about NAO. In Figure 22, the results from the three questions about **the opinion** of the participants **about robots (1)** are shown. This image clearly shows that participants mainly think that robots are not scary. In addition, they view robots as not particularly thrilling/exciting. It is important to mention that the Dutch word that was used in the questionnaire was “Spannend”. This word translates to

both thrilling and exciting depending on the context. In this context, it could mean both but it seems to have been interpreted more as thrilling as the answers seem to be very negative while the opinion about if the robot is cool is very high.

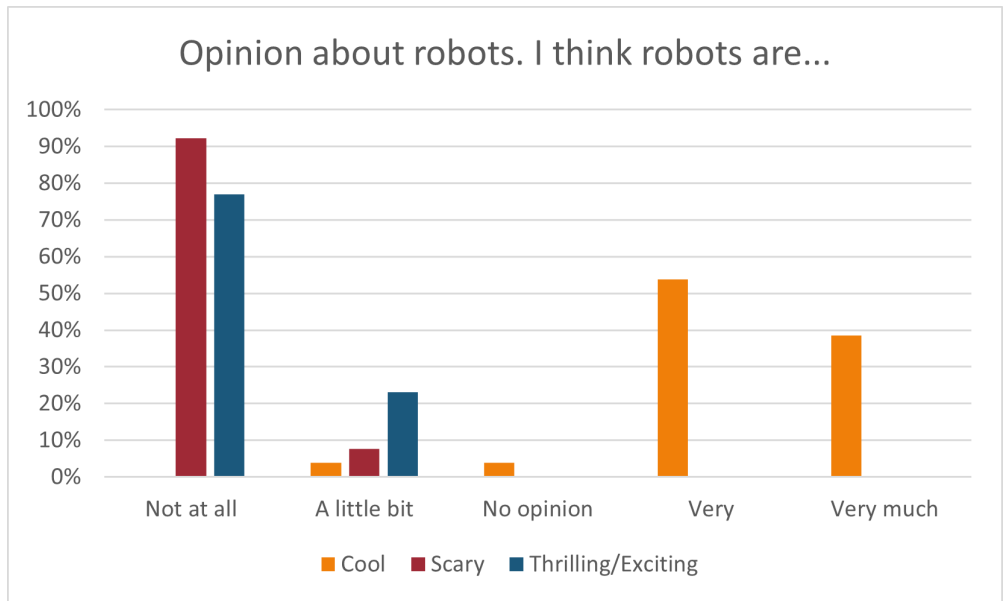


Figure 22: Measured the **(1) opinion about robots** from the participants before the session. The horizontal axis is the agreement with the statement. The vertical axis is the percentage of answers.

In the post-questionnaire, the participants were asked about **the opinion of NAO (6)**. The results from this question can be found in Figure 23. The participants thought that the robot was either not at all or a little bit fake. They were quite neutral about the likeliness to a human. They were mostly positive about how smart and nice/kind the robot was.

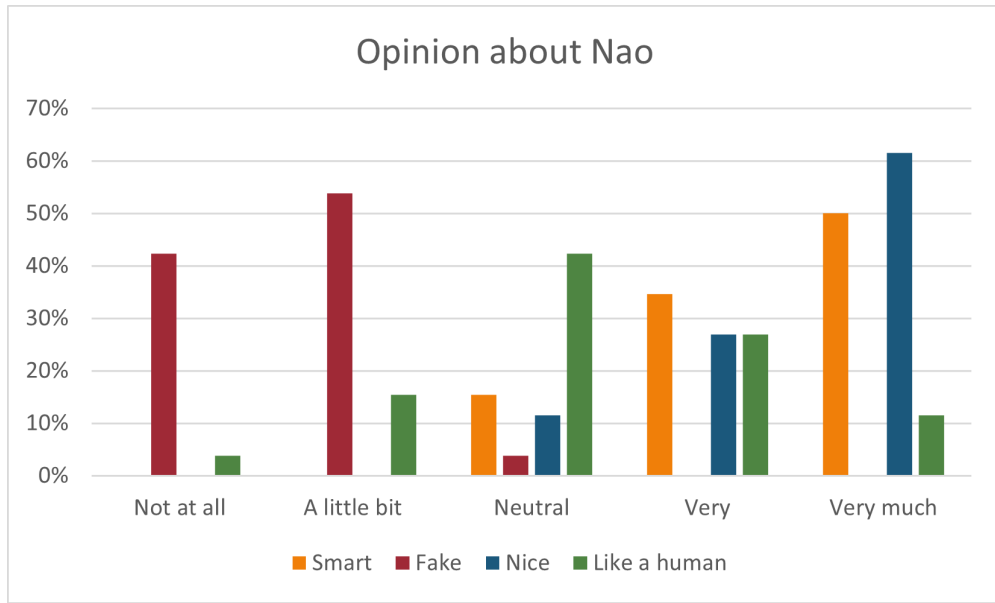


Figure 23: Measured the opinion about the NAO robot from the participants after the session. The horizontal axis is the agreement with the statement. The vertical axis is the percentage of answers.

We make the assumption that due to the participants not being afraid of robots and thinking that robots are cool, they are more willing to talk to a robot than a participant that is afraid of robots. In addition, the participants liked the robot which makes them more likely to talk to the robot in the future.

**Remembering learning goals** The amount of learning goals that the participant remembers partly shows how motivated and active the participants were in learning the goals.

After each level has been completed and the additional questions had been asked by the robot, the participants were asked to recall the learning goals within the game and to reiterate them. The results of how many **learning goals** are **remembered** (4) are shown in Figure 24. As there were twenty-five audio recordings, the maximum amount would be twenty-five. This means that the fourth learning goal only had five students not remembering it. There does not seem a clear rise in the learning goals remembered so we can assume that the time that the learning goals are being recalled does not have an impact on the number of remembered learning goals.



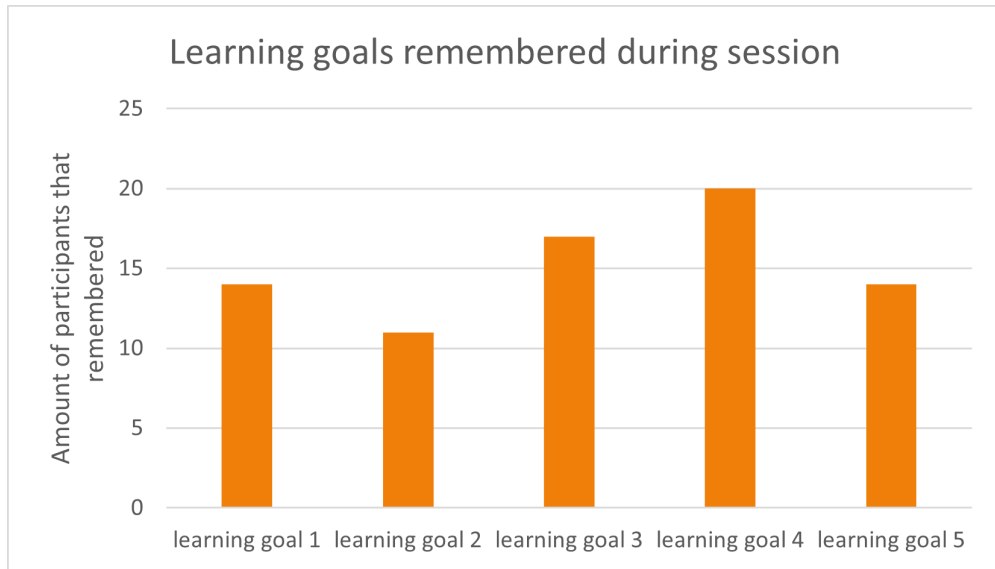


Figure 24: Measured which learning goals have been remembered in-between levels. The horizontal axis represents the five learning goals. The vertical axis represents the number of participants that remembered the learning goal. As there were twenty-five participants, the maximum amount would be twenty-five.

In the post-session questionnaire, the participants filled in the questionnaire described in subsection 5.4.3. The entire questionnaire can be found in subsection A.2.

In this questionnaire, the participants were asked to summarise what they learned from the levels. Some participants understood that this meant the learning goals while others either did not or did not remember the learning goals. The **number of learning goals that have been remembered (7)** can be found in Figure 25. There were twenty-five answers that were given so the max in this figure would be twenty-five. Learning goal three was remembered best as ten out of twenty-five participants remembered this learning goal.

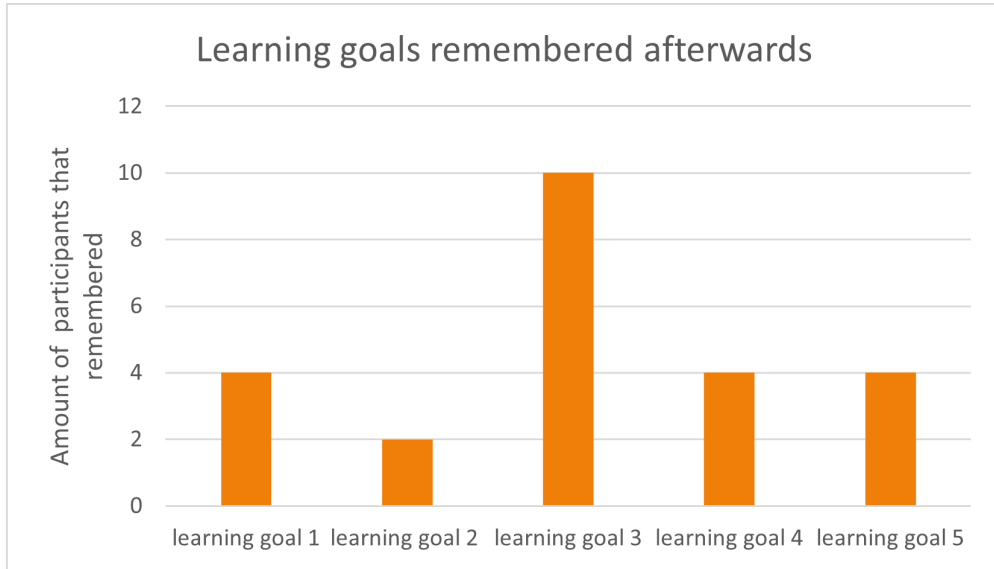


Figure 25: Measured which learning goals have been remembered after the session. The horizontal axis represents the five learning goals. The vertical axis represents the number of participants that remembered the learning goal. As there were twenty-five participants, the maximum amount per learning goal would be twenty-five.

In Figure 26, the number of learning goals that each participant remembered during and after the session is recorded. As is shown, during the session, most participants remembered three goals. After the session, most participants remembered zero learning goals.

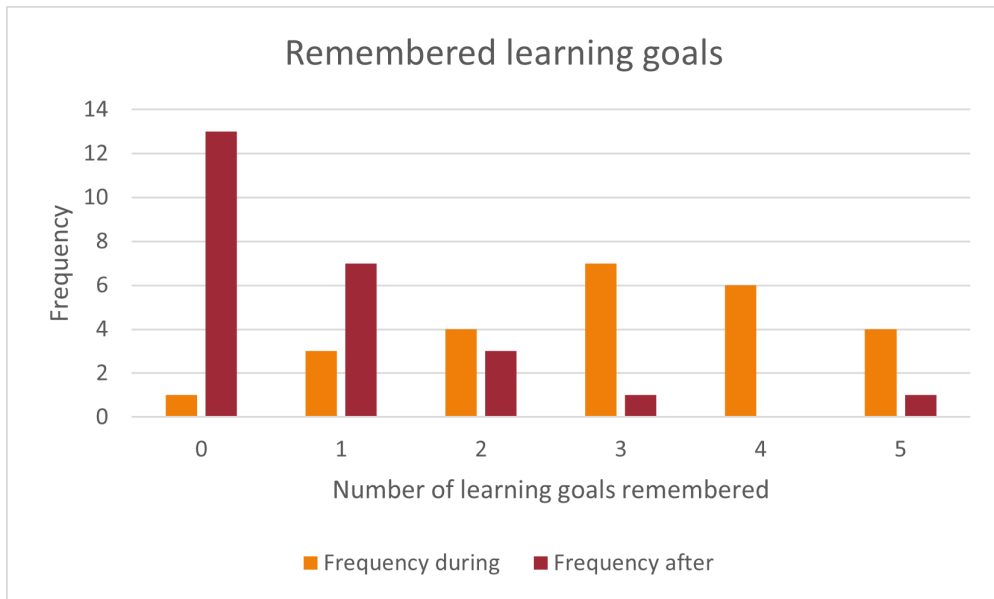


Figure 26: Measured the number of learning goals remembered after the session. The horizontal axis represents the five learning goals. The vertical axis represents the number of participants that remembered the learning goal. As there were twenty-five participants, the maximum amount would be twenty-five.

## 7 Discussion

The outcomes of this research have provided insight into the two different manners of gaining insight into the values of children. The implications of the results are discussed in subsection 7.1. The results of this paper should be interpreted with caution due to the limitations as mentioned in subsection 7.2. Finally, suggestions for further research will be discussed in subsection 7.3.

The robot assisted the participant in remembering the learning goals by ensuring that the participant read the learning goals in the game and helped by repeating the learning goals if the participants forgot the learning goals at the end of the level. The robot assisted with reflecting on decisions and gaining insight into values. Three hypotheses have been set up and have been evaluated about gaining insight into values and the quality of the argumentation connected to the insight.

Hypothesis one a aimed to test whether the method of Clarification as a form of Socratic Questioning by a robot would provide better insight into a child's values compared to the Implications & Consequences type of Socratic questioning. Visually, as can be seen in Figure 4, Clarification had better results except for questions one and five. A factor that questions one and five have in common is that they are the first questions to be asked once a level is finished. Questions nine and thirteen are also the first questions about the level and although the Clarification question does score better, there still is an increase in score for the Implication and Consequences question compared to the surrounding questions. Another factor that they have in common is that the Implication and Consequences questions for numbers one and five are both questions that ask about how the other person in the story would feel about the decision that has been made. Interestingly, questions two, four and fourteen are similar questions but ask about how others will think or respond to the decision. It seems that asking the participants about the feelings of other children instead of the opinion of other people will give a better insight into their own values.

Hypothesis one b tested if a recurrence of a Socratic Question led to a better (value-based) argumentation quality than the previous Socratic Questions. A clear increase is visible in Figure 8. Statistically, the recurrence of the questions is of impact on the quality of argumentation. This means that the more questions are asked, the better the quality of argumentation becomes.

The statistical analysis showed a significant difference between the average scores per question type. Specifically questions two, four, fourteen and sixteen showed a significant difference in scores. All of these differences were in favour of Clarification. Interestingly, the Implication and Consequences questions of questions two, four and fourteen ask about how other children would think or respond to their choice. It does not seem that the Implication and Consequences scores of questions four, fourteen and sixteen are significantly lower than the surrounding Implication and Consequences questions, it is more likely that the Clarification questions were easier to answer with a better quality of argumentation. There is a significantly lower score for the Implications & Consequences question at question two. This is mainly due to a lot of the participants answering the question in a very short manner; they are asked what Jan and Peter would think about their decision and quite some answers were similar to: They will think I made a good decision. This is an answer that shows a bad quality of argumentation which is why the score is quite low.

There is a significant difference between the Clarification and Implication and Consequences questions. As Clarification had a higher average score, it indicates that Clarification leads to a better quality of argumentation.

The second hypothesis focused on examining the interaction effect between the recurrence of the question type and the question type on the quality of argumentation. The Univariate Analysis of

Variance showed a significant effect of the question type and the recurrence. An interaction effect between the question type and recurrence was also prevalent. This shows that the effect of question type on the quality of argumentation varied depending on the recurrence.

Finally, hypothesis three investigated whether the personality traits of the participant had an impact on the quality of argumentation. The stepwise regression showed that Extraversion had the most impact but there was no significant effect on the quality of argumentation.

The qualitative observations mainly showed observations about the actions of the participant. The first observation showed that the participants gave explanations without being prompted by the robot. This indicated that the participants were anticipating the robot's questions. In addition, some participants gave answers that aligned with different values than initially connected to the question. Nearly half (46%) of the responses did not correspond to the self-reported order of value importance. This indicates that our personally held beliefs regarding the order of importance of values are context-dependent.

## 7.1 Implications

There is a very limited amount of research comparing multiple methods of Socratic Questioning. For the dilemmas in this experiment, the Clarification questions lead to the best arguments of children in general. However, the Implications & Consequences questions that asked about the opinion of other children about the decisions of the participant always performed better than the Clarification questions in the same dilemma. This insight should be used in research that further gains insight into the values of children.

Hypothesis three showed that we did not find an effect of personality on the quality of argumentation. In the research done by Abbassi-asl [58], the personality of the participants is compared to the four components of moral decision-making. Moral reasoning is one of these factors. Abbassi-asl describes moral reasoning as "judgment in regards to which action is morally right or not". One's morals are similar to one's values, morals are generally more often influenced by the society around a person while values are more often related to what the person wants themselves. Abbassi-asl uses the term of moral values when describing morals which is why Moral reasoning can be compared to the reasoning based on values that the participants in this study performed. Abbassi-Asl reported that personality traits could not predict moral reasoning. This is in line with the results of our study.

In the qualitative observations, it was noticed that almost half of the time, the values of the participants did not correspond with the self-reported values. This is in line with the results that Hertz reported [59]. Hertz reported that "behavioral observations and self-reported behavioral intentions tended to be less strongly correlated with moral identity than self-reports and other reports of past behavior". Hertz also mentions that this may be caused due to behavioural observations and self-reported intentions being typically tied to specific situations and scenarios.

## 7.2 Limitations

**Question two.** There is an anomaly in the data from question two. The results from this are significantly lower than the surrounding questions. This could be explained by the set-up between questions one and two. In question one, the robot asked why the participant made a decision to choose that specific person to be the DJ. The second question simplified the first question by removing a limiting factor that could alter their decision. Then the robot asked again why they would or would not change their decision. Many answers to this question would be similar to: "I would not change my answer, I like it this way". Which does not show a value and would result in a low score. Some participants did not feel the need to explain their values as they would often be the same as question

one. This resulted in a lower score for the question.

**Small sample size.** The generalizability of the results is limited by the small sample size. The experiment was performed on twenty-six children between nine and eleven years old. To reduce the effects of the small sample size, the tests were performed within-subject. This allows for more data points as all participants experienced all types of questions. In addition, the variance between individuals is better addressed due to the within-subject study.

**Homogeneous sample.** The school is situated in a neighbourhood with a higher proportion of higher-educated residents. This can slightly impact the personality of the participant, the familiarity with robots, and the opinion about education. The quality of argumentation could also be impacted as such a group could have better textual argumentation skills.

**Subjective rating quality of argumentation.** The next limitation is the method of rating the quality of argumentation. Some answers show a very clear insight and would result in a perfect score and some did not show any insight at all and would receive a very low score. However, some answers are quite difficult to rate. These answers were rated based on the interpretation of the researcher. To reduce the subjectivity, the answers were rated based on a list with all answers given to the same question. To further reduce subjectivity, another researcher was asked to give a rating to ten random answers. This researcher gave very similar answers. An interrater reliability test has been done which showed that the agreement is above chance agreement.

**Environmental noise.** As the experiment occurred in a school, some environmental noise was expected. The school has split the break where the students play outside into multiple parts; one for the lowest grades, one for the middle grades and one for the highest grades. This resulted in some noise during the session by children running in the hallway or screaming while playing outside. The participants were used to this noise and did not seem to be bothered at all by this noise. During one session, the classroom next door started playing Just Dance, this resulted in some very loud music which could be heard clearly in the room. The participant did not seem bothered by the music but the researchers did have some difficulty in understanding the child in two answers. In one answer, this did result in the wrong Wizard of Oz button being pressed which resulted in the robot asking the participant to repeat the learning goal while they just said it. The researcher quickly pressed the button again, and the robot would complement the student on remembering the learning goal. The incident was explained as the robot having some trouble hearing the student due to the noise. After three songs, the noise stopped. Finally, during the experiment of participant number 25, another student interrupted and asked an important question to the participant. The participant answered and the other student left the room again. The participant did not seem bothered by the interruption and had no trouble recollecting the learning goal.

**Robot falling over.** During the session with participant number 3, the robot fell over in the middle of a sentence. The robot used seemed to have an issue where the robot would suddenly collapse as if turned off. Luckily this only happened twice, once during set-up and once during the session with the participant. The robot had just been connected to the charger about a minute ago and the researcher told the child that the robot probably tripped over their charger as the participant feared that they caused the robot to fall. The participant was reassured that it was not their fault. The participant continued the session without any further incident.

**Robot lost connection.** In five sessions, the robot lost connection in the middle of the session. In this case, the researcher would remove the code that had already been said and would restart the robot where it left off once the connection was restored.

**Difficulties understanding the robot.** In almost all sessions, the participants had some trouble understanding the robot. To start, the robot has a weird pronunciation of some words. To solve this, some words were coded with spelling mistakes to ensure that the robot was able to pronounce them slightly better but unfortunately, not all words were able to be pronounced by the robot. In addition, the pacing of the robot was not correct. Sometimes, the robot would talk really fast, skipping all punctuation marks and sometimes it would stop for a bit longer at punctuation marks. This issue caused a lot of the participants to either give weird answers, to not answer at all or to ask the researcher what the robot said. In all of these scenarios, the researcher asked the participant if they understood the question and mentioned that the robot often talks weirdly and that it is not their fault that they did not understand it. After this, they would repeat the question to the participant if necessary.

**Answers before the question.** During the experiment, some participants were keen to answer the initial question and give an explanation without being prompted by the robot. For example, the robot would ask: “Would you choose Jan or Peter?”. An example of an answer from a participant would be: “I would choose Jan because he was really enthusiastic and Peter was not”. Then the robot would ask the Clarification question: “Why?” and the participant would sometimes repeat their reason and sometimes say that they have already answered the question. This sometimes resulted in answers being less insightful in values.

**Clicking done before the robot could say so.** As explained in subsection 3.2, the game tells the player that the robot will now talk to them and to not press the button to continue unless the robot asks them to continue. Some participants clicked the button before the robot said that they could. This often distracted the player as they did not listen to the robot and wanted to continue the game. The researchers had to ask the participant to stop playing the game and listen to the robot.

## 7.3 Further Research

Some suggestions for further research will be given in this section.

### 7.3.1 Other types of Socratic Questioning

As discussed in subsection 4.1, only Clarification and Implications & Consequences are compared for the insight into values of children. There are four other types of Socratic Questioning that could be compared: Questions that probe assumptions, questions that probe reason and evidence, Questions that probe viewpoints and perspective, and Questions that probe questions [12]. These types of Socratic Questioning are considered as quite difficult to answer. By comparing these types of questions to gain insight into values, it may be possible to identify the most effective approach for gaining such insight.

### 7.3.2 Better way of rating answers

As mentioned before, the rating of the answers is subjective, even though there were some steps taken to make the rating system more objective. Further research could benefit from having a team of specialists rate the answers and combine their ratings into a single score.

### 7.3.3 Personality with a bigger sample size

The results to see if personality had a significant influence on the insight into values, a regression test has been done. The regression test showed a probability that was really close to being significant.

Further research could perform a similar test with a bigger and more diverse participant group to see if personality has any influence.

### **7.3.4 Add additional types of neurodiversity**

The game is currently only showing learning goals about ADHD, ASD and Dyslexia. Considering the limited amount of teaching about neurodiversity and the existence of more types of neurodiversity, conducting further research to incorporate more types and create a more inclusive game would be beneficial.

### **7.3.5 Test the true influence of a game on the motivation**

As mentioned before in subsection 6.5, the participants did not complain about the number of questions that they had to answer compared to the previous experience of the robot camp. However, there are two major factors that could have caused this influence: the setting and the game. This experiment was done during school time when the participants were missing class compared to the robot camp where the participants were missing time in which they could play with a robot. The other difference is that the children in the robot camp were given initially eight scenarios with pictures where they had to make a decision. The children in this experiment were given a level of a game where one scenario occurred and then a question about that scenario was given to them, after that, they were given three scenarios with questions. In total the participants did four levels which resulted in sixteen questions. It is suspected that the game has a major influence on the motivation of the participants to answer the questions to gain insight into values. However, due to there being two major possible influences, future work could check if the game has a major influence on increasing motivation. Future work could give the same questions with and without a game and could see if the game increases the motivation, or reduce the boredom, of the participants.

## **7.4 Recommendations**

### **7.4.1 Explanation at the choice question**

When testing different methods of questioning, it is recommended to ensure that participants are not giving an explanation at the dilemma question as it results in data that can only partly be used. To prevent participants from giving an explanation when their choice is asked, two possible methods are suggested. The first method is interrupting the participant when they are giving an explanation. The robot could ask them to wait with giving further explanation. The other method is presenting the participants with a structured set of response options. This forces them to choose from the provided answers rather than offering open-ended explanations. Limiting this set of options could also reduce the “swaps to another value”. The participants are more limited in their options and would be less likely to answer with a value that does not correspond with the values connected to the dilemma. However, this method removes some freedom of creativity in the experiment.

## **8 Conclusions**

In this study, the robot played a role in helping participants remember the learning goals by ensuring they read them when playing the game. In addition, the robot provided assistance by reiterating the learning goals if the participants forgot them at the end of each level. The robot also supported the participants in reflecting on their decisions and gaining insight into their values.

This research aimed to contribute by analyzing multiple methods of gaining insight into the values of the participant. The research also analyzed if the participants’ personality traits influenced the gained insight into values. In addition, this research contributed by delivering an educational game

on ADHD, ASD and Dyslexia.

Based on the findings of the study, several conclusions can be drawn. First, the results indicate that Clarification as a form of Socratic Questioning by a robot provides a better insight into a child's values than the Implications & Consequences type of Socratic questioning. In dilemmas where participants were asked to consider how other children would feel about the decision made by the participant, the results showed better results when using an Implications & Consequences question over a Clarification. This shows a possible advantage of Implications & Consequences if used in this manner.

Second, there was a learning effect for both types of questions. The quality of the argumentation on the Implication & Consequences question was lower at the start but seemed to have a steeper learning curve, showing a significant interaction effect. This type of question could be useful when question-evoked argumentation experience increases.

Finally, the impact of personality traits on the quality of argumentation was investigated. This showed that, while Extraversion had the most influence, it did not show a significant effect on the quality of argumentation.

Next to the quantitative findings, the qualitative observations provided some additional valuable insights. Participants had the tendency to give an explanation of why a decision has been made before the robot prompted an explanation. This indicated some anticipation of the robot's questions. Some participants gave answers that aligned with other values than the values that were connected with the choices. Finally, almost half of the responses did not align with the participants' self-reported order of value importance. This suggests that context influences decisions so they diverge from the self-reported general priority of values.

In conclusion, this study shows that Clarification results in a better insight into the values of children. The interaction between the question type and the recurrence further shows the importance of timing during questioning. These findings contribute to understanding how to effectively gain insight into values in an educational and research setting.

## References

- [1] *Cijfers over basisscholen*. [Online]. Available: <https://www.nji.nl/cijfers/basisscholen>.
- [2] D. Ince, *Cijfers over pesten: Slachtoffers*. [Online]. Available: <https://www.nji.nl/cijfers/pesten-slachtoffers>.
- [3] G. Vessels and W. Huitt, "Moral and character development," *the National Youth at Risk Conference*, 2005.
- [4] M. Gervais, "Exploring moral values with young adolescents through process drama," *International Journal of Education & the Arts*, vol. 7, no. 2, Apr. 2006.
- [5] O. T. ten Cate, R. A. Kusurkar, and G. C. Williams, "How self-determination theory can assist our understanding of the teaching and learning processes in medical education. amee guide no. 59," *Medical Teacher*, vol. 33, no. 12, pp. 961–973, 2011. DOI: 10.3109/0142159x.2011.595435.
- [6] NOS, *Lerarentekort op basisscholen gestegen tot bijna 10 procent*, Dec. 2022. [Online]. Available: <https://nos.nl/artikel/2456228-lerarentekort-op-basisscholen-gestegen-tot-bijna-10-procent>.



- [7] M. van Algemene Zaken, *Hoe is de groep van mijn kind op de basisschool samengesteld?* Dec. 2021. [Online]. Available: <https://www.rijksoverheid.nl/onderwerpen/basisonderwijs/vraag-en-antwoord/ho-zijn-de-groepen-in-het-basisonderwijs-bo-samengesteld#:~:text=Gemiddelde%5C%20groeps-grootte%5C%20basisonderwijs,en%5C%20in%5C%202020%5C%2022%5C%20C9..>
- [8] E. Endlich, *Neurodiversity in college admissions*, Aug. 2022. [Online]. Available: <https://www.iecaonline.com/quick-links/ieca-news-center/insights-newsletter/insights-summer-2022/neurodiversity-in-college-admissions/#:~:text=Neurodiversity%5C%20refers%5C%20to%5C%20the%5C%20fact,patterns%5C%20outside%5C%20the%5C%20neurotypical%5C%20mainstream..>
- [9] V. Rosanda and A. Istenic Starcic, "The robot in the classroom: A review of a robot role," *Emerging Technologies for Education*, pp. 347–357, 2020. DOI: 10.1007/978-3-030-38778-5\_38.
- [10] O. Mubin, C. J. Stevens, S. Shahid, A. A. Mahmud, and J.-J. Dong, "A review of the applicability of robots in education," *Technology for Education and Learning*, vol. 1, no. 1, 2013. DOI: 10.2316/journal.209.2013.1.209-0015.
- [11] B. Scassellati, L. Boccanfuso, C.-M. Huang, *et al.*, "Improving social skills in children with asd using a long-term, in-home social robot," *Science Robotics*, vol. 3, no. 21, 2018. DOI: 10.1126/scirobotics.aat7544.
- [12] Y. H. Manurung and F. S. Siregar, "Developing students critical thinking on speaking through socratic questioning method," in *Proceedings of the International Conference on Language, Literature, and Education (ICLLE 2018)*, Atlantis Press, 2018/12, pp. 212–216, ISBN: 978-94-6252-625-9. DOI: 10.2991/iclle-18.2018.35. [Online]. Available: <https://doi.org/10.2991/iclle-18.2018.35>.
- [13] C. A. Padesky, "Socratic questioning: Changing minds or guiding discovery," in *A keynote address delivered at the European Congress of Behavioural and Cognitive Therapies, London*, vol. 24, 1993.
- [14] *Education terms glossary*. [Online]. Available: <https://www.proctorfree.com/education-terms-glossary/>.
- [15] C. C. Abt, *Serious games*. Viking Press, 1970.
- [16] P. J. Shah, M. Boilson, M. Rutherford, *et al.*, "Neurodevelopmental disorders and neurodiversity: Definition of terms from scotland's national autism implementation team," *The British Journal of Psychiatry*, vol. 221, no. 3, pp. 577–579, 2022. DOI: 10.1192/bjp.2022.43.
- [17] M. Legault, J.-N. Bourdon, and P. Poirier, "From neurodiversity to neurodivergence: The role of epistemic and cognitive marginalization," *Synthese*, vol. 199, no. 5-6, pp. 12 843–12 868, 2021. DOI: 10.1007/s11229-021-03356-5.
- [18] C. Ariane Resnick, *What is neurodivergence and what does it mean to be neurodivergent?* Jan. 2023. [Online]. Available: <https://www.verywellmind.com/what-is-neurodivergence-and-what-does-it-mean-to-be-neurodivergent-5196627>.
- [19] *Attention-deficit/hyperactivity disorder*. [Online]. Available: <https://www.nimh.nih.gov/health/topics/attention-deficit-hyperactivity-disorder-adhd>.
- [20] *Autism spectrum disorder*. [Online]. Available: <https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd>.
- [21] *Dyslexia*. [Online]. Available: <https://www.ninds.nih.gov/health-information/disorders/dyslexia>.
- [22] J. S. Eccles, C. Midgley, A. Wigfield, *et al.*, "Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families.," *American Psychologist*, vol. 48, no. 2, pp. 90–101, 1993. DOI: 10.1037/0003-066x.48.2.90.

- [23] S. H. Schwartz, "An overview of the schwartz theory of basic values," *Online Readings in Psychology and Culture*, vol. 2, no. 1, 2012. DOI: 10.9707/2307-0919.1116.
- [24] S. Schwartz, "Basic human values: Theory, measurement, and applications," *Revue Francaise de Sociologie*, vol. 47, pp. 929-968+977+981, Oct. 2006.
- [25] S. Schwartz, "Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries," in Dec. 1992, vol. 25, pp. 1-65, ISBN: 9780120152254. DOI: 10.1016/S0065-2601(08)60281-6.
- [26] A. K. Döring, A. Blauensteiner, K. Aryus, L. Drögekamp, and W. Bilsky, "Assessing values at an early age: The picture-based value survey for children (pbvs-c)," *Journal of Personality Assessment*, vol. 92, no. 5, pp. 439-448, 2010, PMID: 20706930. DOI: 10.1080/00223891.2010.497423. eprint: <https://doi.org/10.1080/00223891.2010.497423>. [Online]. Available: <https://doi.org/10.1080/00223891.2010.497423>.
- [27] A. K. Döring, "Assessing children's values: An exploratory study," *Journal of Psychoeducational Assessment*, vol. 28, no. 6, pp. 564-577, 2010. DOI: 10.1177/0734282909357151. eprint: <https://doi.org/10.1177/0734282909357151>. [Online]. Available: <https://doi.org/10.1177/0734282909357151>.
- [28] S. H. Schwartz, G. Melech, A. Lehmann, S. Burgess, M. Harris, and V. Owens, "Extending the cross-cultural validity of the theory of basic human values with a different method of measurement," *Journal of Cross-Cultural Psychology*, vol. 32, no. 5, pp. 519-542, 2001. DOI: 10.1177/0022022101032005001. eprint: <https://doi.org/10.1177/0022022101032005001>. [Online]. Available: <https://doi.org/10.1177/0022022101032005001>.
- [29] D. A. Yudkin, A. P. Gantman, W. Hofmann, and J. Quoidbach, "Binding moral values gain importance in the presence of close others," *Nature Communications*, vol. 12, no. 1, 2021. DOI: 10.1038/s41467-021-22566-6.
- [30] T. A. Carey and R. J. Mullan, "What is socratic questioning?" *Psychotherapy: Theory, Research, Practice, Training*, vol. 41, no. 3, pp. 217-226, 2004. DOI: 10.1037/0033-3204.41.3.217.
- [31] J. C. Overholser, "Facilitating autonomy in passive-dependent persons: An integrative model," *Journal of Contemporary Psychotherapy*, vol. 17, no. 4, pp. 250-269, 1987. DOI: 10.1007/bf00949025.
- [32] R. D. Chessick, "Socrates: First psychotherapist," *The American Journal of Psychoanalysis*, vol. 42, no. 1, pp. 71-83, 1982. DOI: 10.1007/bf01253435.
- [33] C. Torabizadeh, L. Hodayuni, and M. Moattari, "Impacts of socratic questioning on moral reasoning of nursing students," *Nursing Ethics*, vol. 25, no. 2, pp. 174-185, 2018, PMID: 27694549. DOI: 10.1177/0969733016667775. eprint: <https://doi.org/10.1177/0969733016667775>. [Online]. Available: <https://doi.org/10.1177/0969733016667775>.
- [34] P. Jeremy Sutton, *Socratic questioning in psychology: Examples and techniques*, Mar. 2023. [Online]. Available: <https://positivepsychology.com/socratic-questioning/>.
- [35] A. Edwards, C. Edwards, P. R. Spence, C. Harris, and A. Gambino, "Robots in the classroom: Differences in students' perceptions of credibility and learning between "teacher as robot" and "robot as teacher"," *Computers in Human Behavior*, vol. 65, pp. 627-634, 2016, ISSN: 0747-5632. DOI: <https://doi.org/10.1016/j.chb.2016.06.005>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0747563216304332>.
- [36] F. Tanaka and S. Matsuzoe, "Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning," *Journal of Human-Robot Interaction*, pp. 78-95, 2012. DOI: 10.5898/jhri.1.1.tanaka.

- [37] O. A. B. Henkemans, B. P. Bierman, J. Janssen, *et al.*, “Design and evaluation of a personal robot playing a self-management education game with children with diabetes type 1,” *International Journal of Human-Computer Studies*, vol. 106, pp. 63–76, 2017, ISSN: 1071-5819. DOI: <https://doi.org/10.1016/j.ijhcs.2017.06.001>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1071581917300861>.
- [38] R. Peters, J. Broekens, and M. A. Neerincx, “Guidelines for tree-based collaborative goal setting,” *Proceedings of the 22nd International Conference on Intelligent User Interfaces*, 2017. DOI: 10.1145/3025171.3025188.
- [39] M. van den Berg, *Learning to eat healthily with a robot*, Jul. 2022.
- [40] L. E. Boyd, K. Day, N. Stewart, K. Abdo, K. Lamkin, and E. Linstead, “Leveling the playing field: Supporting neurodiversity via virtual realities,” English, *Technology and Innovation*, vol. 20, no. 1, pp. 105–116, 2018, Copyright - Copyright National Academy of Inventors 2018; Last updated - 2022-11-06. [Online]. Available: <https://www.proquest.com/scholarly-journals/leveling-playing-field-supporting-neurodiversity/docview/2174212450/se-2>.
- [41] S. p. Santhanam, “An interactive and neurodiversity-affirming approach to communication supports for autistic students through videogaming,” *Language Speech and Hearing Services in Schools*, Jan. 2023. DOI: 10.1044/2022\_LSHSS-22-00027.
- [42] K. Young, 2018. [Online]. Available: <https://www.artstation.com/artwork/mk1P8>.
- [43] H. López-Carral, M. Blancas-Muñoz, A. Mura, *et al.*, “A virtual reality system for the simulation of neurodiversity,” *Proceedings of Sixth International Congress on Information and Communication Technology*, pp. 523–531, 2021. DOI: 10.1007/978-981-16-2380-6\_46.
- [44] S. Ledger, “On dis/ability within game studies,” *Interdisciplinary approaches to disability: Looking towards the future*, vol. 2, 2018.
- [45] E. Wadlington, C. Elliot, and J. Kirylo, “The dyslexia simulation: Impact and implications,” *Literacy Research and Instruction*, vol. 47, no. 4, pp. 264–272, 2008. DOI: 10.1080/19388070802300363. eprint: <https://doi.org/10.1080/19388070802300363>. [Online]. Available: <https://doi.org/10.1080/19388070802300363>.
- [46] Z. Stark, L. Franzen, and A. P. Johnson, “Insights from a dyslexia simulation font: Can we simulate reading struggles of individuals with dyslexia?” *Dyslexia*, vol. 28, no. 2, pp. 228–243, 2022. DOI: <https://doi-org.tudelft.idm.oclc.org/10.1002/dys.1704>. eprint: <https://onlinelibrary-wiley-com.tudelft.idm.oclc.org/doi/pdf/10.1002/dys.1704>. [Online]. Available: <https://onlinelibrary-wiley-com.tudelft.idm.oclc.org/doi/abs/10.1002/dys.1704>.
- [47] L. E. Meinen, “Share the experience, don’t take it: Toward attunement with neurodiversity in videogames,” *Games and Culture*, p. 15 554 120 221 149 538, DOI: 10.1177/15554120221149538. eprint: <https://doi.org/10.1177/15554120221149538>. [Online]. Available: <https://doi.org/10.1177/15554120221149538>.
- [48] B. Scassellati, H. Admoni, and M. Mataric, “Robots for use in autism research,” *Annual Review of Biomedical Engineering*, vol. 14, no. 1, pp. 275–294, 2012, PMID: 22577778. DOI: 10.1146/annurev-bioeng-071811-150036. eprint: <https://doi.org/10.1146/annurev-bioeng-071811-150036>. [Online]. Available: <https://doi.org/10.1146/annurev-bioeng-071811-150036>.
- [49] S. Turkay and C. K. Kinzer, “The effects of avatar-based customization on player identification,” in *Gamification: Concepts, methodologies, tools, and applications*, IGI Global, 2015, pp. 247–272.
- [50] S. Trepte and L. Reinecke, “Avatar creation and video game enjoyment,” *Journal of Media Psychology*, vol. 22, no. 4, pp. 171–184, 2010. DOI: 10.1027/1864-1105/a000022.

- [51] A. Bateman, “Responding to children’s answers: Questions embedded in the social context of early childhood education,” *Early Years*, vol. 33, no. 3, pp. 275–288, 2013. DOI: 10.1080/09575146.2013.800844. eprint: <https://doi.org/10.1080/09575146.2013.800844>. [Online]. Available: <https://doi.org/10.1080/09575146.2013.800844>.
- [52] H. Cameron, “Asking the tough questions: A guide to ethical practices in interviewing young children,” *Early Child Development and Care*, vol. 175, no. 6, pp. 597–610, 2005. DOI: 10.1080/03004430500131387. eprint: <https://doi.org/10.1080/03004430500131387>. [Online]. Available: <https://doi.org/10.1080/03004430500131387>.
- [53] J. Wilson, M. Powell, and P. Freeman, “A guide to interviewing children: Essential skills for counsellors, police, lawyers and social workers,” *Child & Family Social Work - CHILD FAM SOC WORK*, vol. 7, pp. 330–331, Nov. 2002. DOI: 10.1046/j.1365-2206.2002.t01-7-00258.x.
- [54] J. Arnall, *When do children understand “consequences?”*, Feb. 2019. [Online]. Available: <https://judyarnall.com/2019/02/18/when-do-children-understand-consequences/>.
- [55] M. Maćkiewicz and J. Ciecuch, “Pictorial personality traits questionnaire for children (pptq-c)—a new measure of children’s personality traits,” *Frontiers in Psychology*, vol. 7, 2016. DOI: 10.3389/fpsyg.2016.00498.
- [56] J. L. Fleiss, “Measuring nominal scale agreement among many raters.,” *Psychological Bulletin*, vol. 76, no. 5, pp. 378–382, 1971. DOI: 10.1037/h0031619.
- [57] K. Kiili, S. de Freitas, S. Arnab, and T. Lainema, “The design principles for flow experience in educational games,” *Procedia Computer Science*, vol. 15, pp. 78–91, 2012, 4th International Conference on Games and Virtual Worlds for Serious Applications(VS-GAMES’12), ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2012.10.060>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050912008228>.
- [58] R. Abbasi-Asl and S. Hashemi, “Personality and morality: Role of the big five personality traits in predicting the four components of moral decision making,” Dec. 2019. DOI: 10.31234/osf.io/6azqs.
- [59] S. G. Hertz and T. Krettenauer, “Does moral identity effectively predict moral behavior?: A meta-analysis,” *Review of General Psychology*, vol. 20, no. 2, pp. 129–140, 2016. DOI: 10.1037/gpr0000062. eprint: <https://doi.org/10.1037/gpr0000062>. [Online]. Available: <https://doi.org/10.1037/gpr0000062>.
- [60] *Digital around the world - datareportal – global digital insights*. [Online]. Available: <https://datareportal.com/global-digital-overview>.
- [61] “The benefits of playing video games.,” *American Psychologist*, vol. 69, no. 1, pp. 66–78, Jan. 2014. DOI: 10.1037/a0034857.
- [62] C. Green and D. Bavelier, “Learning, attentional control, and action video games,” *Current Biology*, vol. 22, no. 6, 2012. DOI: 10.1016/j.cub.2012.02.012.
- [63] D. H. Uttal, N. G. Meadow, E. Tipton, *et al.*, “The malleability of spatial skills: A meta-analysis of training studies.,” *Psychological Bulletin*, vol. 139, no. 2, pp. 352–402, 2013. DOI: 10.1037/a0028446.
- [64] L. A. Jackson, E. A. Witt, A. I. Games, H. E. Fitzgerald, A. von Eye, and Y. Zhao, “Information technology use and creativity: Findings from the children and technology project,” *Computers in Human Behavior*, vol. 28, no. 2, pp. 370–376, 2012. DOI: 10.1016/j.chb.2011.10.006.
- [65] C. Russoniello, K. O’Brien, and J. Parks, “Eeg, hrv and psychological correlates while playing bejeweled ii: A randomized controlled study,” *Studies in health technology and informatics*, vol. 144, pp. 189–92, Feb. 2009.
- [66] R. M. Ryan, C. S. Rigby, and A. Przybylski, “The motivational pull of video games: A self-determination theory approach,” *Motivation and Emotion*, vol. 30, no. 4, pp. 344–360, 2006. DOI: 10.1007/s11031-006-9051-8.

- [67] M. Csikszentmihalyi, K. R. Rathunde, S. Whalen, and M. Wong, *Talented teenagers: The roots of success and failure*. Cambridge University Press, 2000.
- [68] L. P. Rieber and D. Noah, "Games, simulations, and visual metaphors in education: Antagonism between enjoyment and learning," *Educational Media International*, vol. 45, no. 2, pp. 77–92, 2008. DOI: 10.1080/09523980802107096. eprint: <https://doi.org/10.1080/09523980802107096>. [Online]. Available: <https://doi.org/10.1080/09523980802107096>.

## A Questionnaires

### A.1 Pre-session

# Voor experiment

\* Vereist



1. De juf of onderzoeker vult hier een code in voor jou \*

2. Ik vind robots

	Helemaal niet	Een heel klein beetje	Geen mening	Erg	Heel erg
Cool	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Eng	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Spannend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>





### 3. Persoonlijkheid

	Nooit	Heel soms	Gemiddeld	Vaak	Heel vaak
Ik speel meestal alleen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Als ik naar school ga ben ik zenuwachtig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik zie vaak vogels vliegen als ik buiten speel	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik doe graag mijn huiswerk	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik merk het als iemand in mijn klas iets nodig heeft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Als anderen plezier hebben, doe ik met ze mee	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Als er iets mis gaat, wordt ik snel zenuwachtig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Op reis houdt ik van ontspannen in plaats van iets nieuws ontdekken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In mijn kamer is het opgeruimd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Als ik iemand zie die ik kan helpen, dan help ik	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Als iemand

Als iemand een grapje maakt, dan zie ik vaak niet waarom het grappig is

Ik maak mij vaak zorgen over iets

Ik vind het leuk om nieuwe en moeilijke dingen te leren

Als ik geld krijg, dan spaar ik het

Als ik een nieuw spelletje heb, dan deel ik het

Als er een moeilijk probleem is, dan los ik het op

Als iemand iets anders wilt dan ik, dan zorg ik dat ik mijn zin krijg.

---

Deze inhoud is niet door Microsoft gemaakt noch goedgekeurd. De gegevens die u verzendt, zal worden gestuurd naar de eigenaar van het formulier.

 Microsoft Forms



## A.2 Post-session

# NAO Robot vragen

\* Vereist

1. De onderzoeker vult hier een code in voor jou \*

De waarde moet een getal zijn

2. Wat heb je geleerd over de kinderen in het eerste verhaaltje? (Dat je een DJ moest uitkiezen) \*

3. Wat heb je geleerd over de kinderen in het tweede verhaaltje? (Dat je een aan de juf een vraag ging stellen) \*

4. Wat heb je geleerd over de kinderen in het derde verhaaltje? (Dat de kinderen ruzie hadden) \*

5. Wat heb je geleerd over de kinderen in het vierde verhaaltje? (Dat je samen ging werken aan een opdracht) \*

6. Wat vond je van de robot \*

	Helemaal niet	Een beetje	Een beetje wel / een beetje niet	Erg	Heel erg
Slim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nep	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aardig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zoals een mens	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

---

Deze inhoud is niet door Microsoft gemaakt noch goedgekeurd. De gegevens die u verzendt, zal worden gestuurd naar de eigenaar van het formulier.

 Microsoft Forms

## B Scripts

### B.1 The game

*Italics:* Things that the characters in the game say

**Bold:** Name of NPC who says something

**Blue:** Playername, to be filled in with name of player

Underlined: Button/Option to be clicked by player

**Green:** The action that the player does

**Orange:** Text of learning goal in game (bottom of screen)

**Pink:** Text in the game in the middle of the screen

#### B.1.1 Situation 1

In the first situation, there are 2 characters: Jan and Peter. Peter has some aspects of ASD. Jan has some aspects of ADHD.

Learning goals: Some people can handle busy situations a bit worse than others. Some people have a bit more energy that they need to expend than others

Goal: Pick a DJ to play music.

#### Script

*English:*

The player starts in the classroom with the teacher. Peter is in the auditorium, Jan is on the basketball field.

**Teacher Evelien:** *“Hello **Playername**, I need some help planning the party, can you help me by any chance? I need help choosing a DJ, I heard that **Peter** is a super good DJ but maybe he doesn’t want to be a DJ. I’ve also heard that **Jan** might want to be a DJ. Could you choose who will be the DJ? **Peter** is in the auditorium and **Jan** is on the basketball court. Come back to me when you have talked to them.”*

Option button: *“Okay!”*

**Action:** The player walks in the school

**Text during action:** Some children can handle busy environments a bit worse than others, they might get overstimulated.

**Action:** While the player walks outside

**Text during action:** Some children have a bit more energy to expend than others.

Conversation with **Peter:**

**Playername:** *“Hi **Peter**, I have been instructed by the teacher to find a DJ. The teacher said that you are really good.”*

**Peter:** *“Hi **Playername**. I can choose music pretty well, but at a party, it is often too busy for me, but I can be a DJ.”*

Conversation with **Jan:**

**Playername:** *“Hi **Jan**, I have been instructed by the teacher to find a DJ. The teacher said that you might want to do it.”*

**Jan:** *“Yesssss! I would love to be a DJ! I can also do the decorations but I would really prefer to be a DJ!”*

**Action:** The player returns to the teacher

**Teacher Evelien:** *Who do you want to make the DJ?*

Option button 1: **Peter**

Option button 2: **Jan**

**Text on the screen:** The robot will now ask you a few questions, click done when the robot says you can press done.

Option button: Done

*Dutch:*

Speler begint in de klas bij de juf. Peter staat in de aula, Jan staat op het basketbalveld.

**Juf Evelien:** *“Hallo **Playername**, ik heb wat hulp nodig met het plannen van het feestje, kan jij mij toevallig helpen? Ik heb hulp nodig met het kiezen van een DJ, ik heb gehoord dat **Peter** een super goede DJ is maar misschien wilt hij geen DJ zijn. Ik heb ook gehoord dat **Jan** misschien wel een DJ wilt zijn. Zou jij willen uitkiezen wie de DJ wordt? **Peter** staat in de aula en **Jan** staat op het basketbalveld. Kom maar bij mij terug als je met ze hebt gepraat.”*

Optie knopje: *“Oké!”*

**Actie:** De speler is aan het lopen in de school

**Tekst tijdens actie:** Sommige kinderen kunnen minder goed tegen een drukke omgeving, ze kunnen overprikkeld raken

**Actie:** De speler is buiten aan het lopen

**Tekst tijdens actie:** Sommige kinderen zijn wat drukker dan andere kinderen

Gesprek met **Peter:**

**Playername:** *“Hoi **Peter**, ik heb van de juf de opdracht gekregen om een DJ te vinden voor het feestje van vandaag. De juf zei dat jij heel goed bent.”*

**Peter:** *“Hoi **Playername**. Ik kan wel goed muziek uitkiezen maar bij een feestje is het vaak te druk voor mij, maar ik kan wel DJ zijn hoor.”*

Gesprek met **Jan:**

**Playername:** *“Hoi **Jan**, ik heb van de juf de opdracht gekregen om een DJ te vinden voor het feestje van vandaag. De juf zei dat jij het leuk zou vinden om DJ te zijn”*

**Jan:** *“Jaaaa! Ik vind het echt super leuk om DJ te zijn!”*

**Actie:** De speler gaat terug naar de juf

**Juf Evelien:** *Dus, wie wordt de DJ?*

Optie knopje 1: **Peter**

Optie knopje 2: **Jan**

**Tekst in beeld:** De robot zal je nu een paar vragen stellen, klik op klaar als de robot zegt dat je op klaar mag drukken.

Optie knopje: Klaar



### B.1.2 Situation 2

A student will approach the player and tell them that they have difficulty with grammar but are afraid to ask the teacher because they feel that their question is stupid. They ask the player if they could ask the question.

Learning goals: Some children are more hesitant to ask something to the teacher than other children. It will take extra effort for some students to consume printed texts independently

Goal: Choose whether or not to ask the question to the teacher as if it was their own question or to tell that it was a question from the other student.

#### Script

*English:*

The player starts at the entrance of the classroom. The **teacher** is at the desk, **Sasha** sits in the seat next to the player's seat. There are other students in the classroom as a lesson is going on. The seat of the player is in the back of the classroom.

**Text on screen:** Go to your seat

**Action:** Go to your seat

**Text during action:** Some children are more hesitant to ask something to the teacher than other children.

**Sasha:** *"Hi **Playername**, can you help me? I have a question for the **teacher** but I am afraid to ask it. I think it is a stupid question. Can you ask the question for me? My question is whether the past tense of teach is teached."*

(Note, question in Dutch is different due to Dutch having different grammatical rules. The question is similar in difficulty)

Option button 1: "Okay"

**Action:** Walk to the teacher

**Text during action:** Some children are more hesitant to ask something to the teacher than other children. **Teacher Eveline:** *"Hi **Playername**, do you want to ask a question?"*

Option button 1: "Yes, I wondered what the past tense of teach is?"

Option button 2: *"It is actually **Sasha**'s question but they were afraid to ask it. The question is what the past tense of teach is."*

**Teacher Eveline:** *"Okay! The past tense of teach is taught."*

Option button: "Thank you"

**Text on the screen:** The robot will now ask you a few questions, click done when the robot says you can press done.

Option button: Done

*Dutch:*

De speler begint bij de ingang van de klas. De **juf** zit aan het bureau. **Sasha** zit in de stoel naast de stoel van de speler.

**Tekst in beeld:** Ga naar je stoel

**Actie:** Ga naar de lege stoel

**Tekst tijdens actie:** Sommige kinderen vinden het enger om iets aan de juf te vragen dan andere kinderen.

**Sasha:** “Hoi **Playername**, kan jij mij misschien helpen? Ik heb een vraag voor de **juf** maar ik durf hem niet te stellen. Ik denk dat het een domme vraag is. Wil jij hem voor mij stellen? Mijn vraag is of het woord loopt met *dt* moet.”

Optie knopje: “Oké”

**Actie:** Loop naar de juf

**Tekst tijdens actie:** Sommige kinderen vinden het enger om iets aan de juf te vragen dan andere kinderen.

**Juf Eveline:** “Hoi **Playername**, wil jij wat vragen?”

Optie knopje 1: “Ja, ik vroeg mij af of het woord loopt met *dt* moet?”

Optie knopje 2: “Het is eigenlijk **Sasha**’s vraag maar die durfde hem niet te stellen. De vraag is of het woord loopt met *dt* moet.”

**Juf Eveline:** “Ah oke, nee het woord loopt hoeft niet met *dt*, het woord eindigt namelijk niet op een *d*.”

Optie knopje: “Bedankt”

**Tekst in beeld:** De robot zal je nu een paar vragen stellen, klik op klaar als de robot zegt dat je op klaar mag drukken.

Optie knopje: Klaar

### B.1.3 Situation 3

Two students are on the playground. One of the kids makes a joke, the other kid does not understand the joke and gets angry.

Learning goals: Some children have more difficulty understanding jokes than other children.

Goal: Choose whether to go to the teacher or to intervene.

#### Script

*English:*

The player starts on the school grounds, a bit further from Henk and Micky. Henk stands towards Micky (and Micky towards Henk). The teacher is in the classroom.

**Action:** Player walks to the conversation

**Text during action:** Some children have more difficulty understanding jokes than other children.

**Henk:** “Yo *momma* is so short, *Spongebob* thought she was *Plankton*”

**Micky:** “Hey that’s rude!”

**Text on the screen:** You see **Micky** and **Henk** get angry and will probably fight. What do you do? You can jump in between them and make sure that they stop or you can get the **teacher**. If you get the teacher, there is a chance that they are too late and the kids will already be fighting.

Option button 1: *Run to the **teacher** and tell them that they are fighting*

Option button 2: *Jump in-between and call for them to stop*

**Text on the screen:** The robot will now ask you a few questions, click done when the robot says you can press done.

Option button: Done

*Dutch:*

Speler begint op het schoolplein iets verder van **Henk** en **Micky** af. Henk staat tegenover **Micky**. De **juf** is in de klas.

**Actie:** Speler loopt naar het gesprek

**Tekst tijdens actie:** Sommige kinderen snappen sommige grappen minder goed dan andere kinderen.

**Henk:** *“Je moeder is zo kort dat SpongeBob dacht dat ze plankton was”*

**Micky:** *“Hey dat is niet leuk!”*

**Tekst in beeld:** Je ziet **Micky** en **Henk** boos worden en gaan misschien wel vechten, wat doe je? Je kan ertussen springen en zorgen dat ze stoppen of je kan de **juf** halen. Als je de **juf** haalt is er een kans dat het al te laat is en dat ze dan al aan het vechten zijn.

Optie knopje 1: *Naar de **juf** toe rennen en zeggen dat ze aan het vechten zijn*

Optie knopje 2: *Ertussen stappen en roepen dat ze moeten stoppen*

**Tekst in beeld:** De robot zal je nu een paar vragen stellen, klik op klaar als de robot zegt dat je op klaar mag drukken.

Optie knopje: Klaar

#### B.1.4 Situation 4

The teacher is making groups for a reading and writing task. The player is grouped with Mike. The player knows that Mike has difficulty reading and writing as fast as the player. The player can decide to do most of the work themselves or to give Mike an equal share of work. Learning goals: Some children have more difficulty reading and writing than others.

Goal: Player chooses to do the main work themselves or spread the tasks equally but possibly getting a worse grade?

Script

*English:*

The player starts on the playground. The player needs to walk to the classroom.

**Action:** Walk to the class

**Text during action:** Some children have to put more effort in reading and writing without mistakes than others.

**Teacher Evelien:** *“I am going to divide you in groups for a reading and writing assignment, **Playername** you will be with **Mike** ”*

**Mike:** *“Hey **Playername**, great to be working with you.”*

**Text on the screen:** You know that Mike has a bit more difficulty reading and writing than others. Will you do most of the work yourself to make sure that the assignment will be done faster or will you split the work evenly?

Option button 1: *Do most of it yourself*

Option button 2: *Share equally*

**Text on the screen:** The robot will now ask you a few questions, click done when the robot says you can press done.

Option button: Done

*Dutch:*

De speler begint op de speelplaats. De speler moet naar de klas lopen.

**Actie:** Loop naar de klas

**Tekst tijdens de actie:** Sommige kinderen moeten meer werk steken in foutloos lezen en schrijven dan andere kinderen.

**Juf Evelien:** *“Ik ga jullie in groepjes verdelen voor een opdracht waar je moet lezen en schrijven. **Playername** jij bent ingedeeld met **Mike**”*

**Mike:** *“Hoi **Playername**, leuk om met je samen te werken.”*

**Tekst op het scherm:** Je weet dat Mike wat meer moeite heeft met lezen en schrijven dan andere kinderen. Ga je het grootste deel van het werk zelf doen zodat het sneller af is of wil je de opdrachten eerlijk verdelen?

Option button 1: *Ik ga het meeste zelf doen*

Option button 2: *Ik ga het eerlijk verdelen*

**Tekst in beeld:** De robot zal je nu een paar vragen stellen, klik op klaar als de robot zegt dat je op klaar mag drukken.

Optie knopje: Klaar

## B.2 The robot

When the word <answer> or <antwoord> is in the script, the robot will wait for an answer.

### B.2.1 Situation 1

In this situation, the child chooses between Peter, who is a good DJ but can get overstimulated in busy situations and does not give preference to being a DJ and Jan who really wants to be a DJ but is worse than Peter.

Script

**During level:**

The robot will instruct the child how to walk, how to use doors and will ask the child to read the learning goals. If the children fail to read the learning goal, the robot will read it for them.

**Question 1:**

The child chooses between someone who will be happy if chosen and someone who is considered the best person for the job.

Underlying values: Benevolence vs Achievement. Choosing Jan is considered as benevolence as

he really wants to do the job. Choosing Peter is considered as the value Achievement as you will get the best possible situation.

*English:*

Text before question: You made a good choice in choosing the DJ!

Clarification: Why did you choose him? <answer>

Implications & Consequences: How do you think Jan and Peter will feel because of your choice? <answer>

Text after question: Ah, I understand now why you chose him

*Dutch:*

Tekst voor vraag: Jij hebt een goede keuze gemaakt bij het kiezen van de DJ!

Clarificatie: Waarom heb je hem gekozen? <antwoord>

Implicaties en Consequenties: Hoe denk je dat Jan en Peter zich zullen voelen door je keuze? <antwoord>

Tekst na vraag: Ah ik snap nu waarom jij hem hebt gekozen, dank je wel!

### **Question 2:**

The situation is the same as question 1 but in this case, the hurdle of the best person not liking busy situations is removed.

Underlying values: Similar to question 1 but it takes away a hurdle.

*English:*

Text before question: Would your choice be different if Peter would not have told you that he did not like busy situations? <answer>

Clarification: Why? <answer>

Implications & Consequences: What would Jan and Peter think about your decision? <answer>

Text after question: Thank you for your answer!

*Dutch:*

Tekst voor vraag: Zou je keuze anders zijn geweest als Peter niet had gezegd dat hij drukke situaties niet zo leuk vindt? <antwoord>

Clarificatie: Waarom? <antwoord>

Implicaties en Consequenties: Hoe zouden Jan en Peter over jouw keuze denken? <antwoord>

Tekst na vraag: Dank je wel voor je antwoord!

### **Question 3:**

The child can now choose if they are going to make a choice which will result in someone getting

angry or choose not to make a choice.

Underlying values: Power vs Security. Having the possibility to make a choice is considered as the value Power. Knowing that you will make someone angry might risk your own security and stability of your friendship.

*English:*

Text before question: You know before you make the choice that the other person is going to get angry anyway. Would you still make a choice or would you go to the teacher and tell her to choose for herself <answer>

Clarification: Why did you choose this? <answer>

Implications & Consequences: What would happen if you do this? <answer>

Text after question: Yes, I understand why you choose this!

*Dutch:*

Tekst voor vraag: Je weet voor je de keuze maakt dat de andere persoon sowieso boos gaat worden. Zou je dan nogsteeds een keuze maken of zou je naar de juf stappen en zeggen dat ze zelf moet kiezen? <antwoord>

Clarificatie: Waarom? <antwoord>

Implicaties en Consequenties: Wat zou er gebeuren als je dit doet? <antwoord>

Tekst na vraag: Ja, dat snap ik!

#### **Question 4:**

In this situation, the child is friends with one of the persons and the other person is considered best for the job. The dilemma is choosing between a friend and the best person for the job.

Underlying values: Benevolence vs Achievement. Choosing the friend is positive for Benevolence while choosing the best person is best for Achievement.

*English:*

Text before question: Suppose you are friends with Jan but you know that Peter is a much better DJ and wants to be a DJ just as much as Jan, would you choose your friend or the one who is better? <answer>

Clarification: Why did you choose this? <answer>

Implications & Consequences: How would Jan and Peter respond to this choice? <answer>

Text after question: I understand you a little better now!

*Dutch:*

Tekst voor vraag: Stel je bent bevriend met Jan maar je weet dat Peter veel beter is en het even graag wilt doen als Jan, zou je dan je vriend kiezen of degene die beter is? <antwoord>

Clarificatie: Waarom zou je dit kiezen? <antwoord>

Implicaties en Consequenties: Hoe zouden Jan en Peter reageren op deze keuze? <antwoord>

Tekst na vraag: Ik begrijp je nu een beetje beter!

**End:**

*English:*

Great, we are done with the first level, do you remember the two learning goals?

Yes: Wow, you are doing great, can you tell them to me?

Good answer: Wow, you do not even need me!

Wrong answer: It is going great! I remembered them as well, I will tell them to you. The first is: Some children can handle busy environments a bit worse than others, they might get overstimulated

The second is: Some children have a bit more energy to expend than others.

No/Forgot: Okay, no problem, we are learning together and I remembered them as well, I will repeat them: The first is: Some children can handle busy environments a bit worse than others, they might get overstimulated

The second is: Some children have a bit more energy to expend than others

Good answer: Wow, you do not even need me!

We are now done with the first couple of questions, you can click on done in the game.

*Dutch:*

Super, we zijn nu klaar met het eerste level, weet jij de twee leerdoelen nog?

Ja: Wow, wat goed van jou, kan jij ze voor mij opnoemen?

Goed antwoord: Wow, jij hebt mij niet eens nodig

Verkeerd antwoord: Het gaat al super goed, ik heb ze ook onthouden, ik zal ze nog even aan je vertellen. De eerste is; Sommige kinderen kunnen minder goed tegen een drukke omgeving, ze kunnen overprikkeld raken. De tweede is: Sommige kinderen zijn wat drukker dan andere kinderen

Nee/Vergeten: Oke, geen probleem, we leren samen en ik heb ze ook onthouden, dan herhaal ik ze wel even. De eerste is: Sommige kinderen kunnen minder goed tegen een drukke omgeving, ze kunnen overprikkeld raken. De tweede is; Sommige kinderen zijn wat drukker dan andere kinderen.

Goed antwoord: Wow, jij hebt mij niet eens nodig!

We zijn nu klaar met de eerste paar vragen, je mag op klaar klikken in het spel.

### B.2.2 Situation 2

Another student asks the child if they can ask a question for them as they are afraid to ask it.

Script

**During level:**

The robot will ask the child to read the learning goal. If the children fail to read the learning goal, the robot will read it for them.

**Question 1:**

Another student asks for help in asking a “stupid” question in asking the teacher for them.

Underlying values: Benevolence vs Self-preservation/achievement. Helping the other student is considered benevolent but it can make them look stupid which is why the opposite side is self-preservation/achievement.

*English:*

Text before the question: You helped Sasha, that’s nice of you

Clarification: Why did you choose this option with the teacher? <answer>

Implications & Consequences: How would Sasha feel about your choice? <answer>

Text after question: Thank you, that makes sense!

*Dutch:*

Tekst voor vraag: Je hebt Sasha geholpen, dat is aardig!

Clarificatie: Waarom heb je deze optie bij de juf gekozen? <antwoord>

Implicaties en Consequenties: Hoe zou Sasha zich voelen door jouw keuze?? <antwoord>

Tekst na vraag: Dank je, dat is logisch!

**Question 2:**

In this situation, the question is considered difficult and the answer is considered unknown. The other student still asks the child to ask the question for them.

Underlying values: The same as question 1 but now we take away the fact that the question is considered “stupid” by Sasha.

*English:*

Text before question: If the question would have been something to which you do not know the answer, would you still choose the same option with the teacher? <answer>

Clarification: Why? <answer>

Implications & Consequences: Would the consequences of your choice be different now? <answer>

Text after question: Great answer, thank you!



*Dutch:*

Tekst voor vraag: Als de vraag iets zou zijn waar je het antwoord niet op weet, zou je dan toch dezelfde optie kiezen bij de docent? <antwoord>

Clarificatie: Waarom? <antwoord>

Implicaties en consequenties: Zouden de gevolgen van je keuze nu anders zijn? <antwoord>

Tekst na vraag: Goed antwoord, dank je!

**Question 3:**

In this question, the student can ask the teacher to make exercises more difficult or not ask the teacher.

Underlying values: Tradition vs Stimulation. Not asking the teacher conforms with tradition as they will follow the lesson plan. Asking the teacher to make the lesson more difficult corresponds with the needs for stimulation.

*English:*

Text before question: You believe that the lessons are a bit easy but you know others, like Sasha struggle. Are you going to ask the teacher to make the exercises a bit more difficult so you learn more or not? <answer>

Clarification: Why did you choose this? <answer>

Implications & Consequences: What would the consequences be for the other children? <answer>

Text after question: Great!

*Dutch:*

Tekst voor vraag: Je denkt dat de lessen een beetje makkelijk zijn, maar je weet dat anderen, zoals Sasha, de les nog moeilijk vinden. Ga je de juf vragen om de oefeningen wat moeilijker te maken zodat je meer leert of juist niet? <antwoord>

Clarificatie: Waarom heb je hiervoor gekozen? <antwoord>

Implicaties en consequenties: Wat zouden de gevolgen zijn voor de andere kinderen? <antwoord>

Tekst na vraag: Geweldig, dank je wel!

**Question 4:**

In this question, the student can choose to do extra work for the next week or to play around.

Underlying values: Hedonism vs Conformity. Choosing to not to homework is hedonism as it makes the student happy. When they choose not to play the game but to learn they restrain themselves.

*English:*

Text before question: You are already done with your homework and can make a drawing on your paper but you can also do some extra homework for next week. What do you do? <answer>

Clarification: Why are you going to do this? <answer>

Implications & Consequences: What would be the consequence of both options? <answer>

Text after question: Great answer!

*Dutch:*

Tekst voor vraag: Je bent al klaar met je huiswerk en kunt een tekening gaan maken maar je kunt ook nog wat extra huiswerk maken voor volgende week. Wat doe je? <antwoord>

Clarificatie: Waarom ga je dit doen? <antwoord>

Implicaties en consequenties: Wat zou het gevolg zijn van beide opties? <antwoord>

Tekst na vraag: Goed antwoord!

**End:**

*English:*

Great, we are now done with the second level, do you still remember the learning goal?

Yes: Wow, you are doing great! Can you explain it to me?

Good answer: Wow, you can learn very well!

Wrong answer: It is going great already, I remembered it as well, I will tell it to you. Some children are more hesitant to ask something to the teacher than other children.

No/Forgot: Okay, no problem, we are learning together and I remembered it as well. I will repeat it: Some children are more hesitant to ask something to the teacher than other children.

Good answer: Wow, you are great at learning!

We are now done with the second set of questions, you can click on Done in the game.

*Dutch:*

Super, we zijn nu klaar met het tweede level, weet het leerdoel nog?

Ja: Wow, wat goed van jou, kan jij het even aan mij uitleggen?

Goed antwoord: Wow, jij kan goed leren!

Verkeerd antwoord: Het gaat al super goed, ik heb hem ook onthouden, ik zal hem nog even aan je vertellen. Sommige kinderen vinden het enger om iets aan de juf te vragen dan andere kinderen

Nee/Vergeten: Oke, geen probleem, we leren samen en ik heb hem ook onthouden, dan herhaal ik hem wel even. Sommige kinderen vinden het enger om iets aan de juf te vragen dan andere kinderen.

Goed antwoord: Wow, jij kan echt super goed dingen leren!

We zijn nu klaar met de tweede paar vragen, je mag op klaar klikken in het spel.

### B.2.3 Situation 3

In this situation two kids are in a quarrel because one made a joke about the others mother and the other kid did not understand that it was a joke. The child can choose to interfere or to get a teacher

## Script

### **During level:**

The robot will ask the child to read the learning goal. If the children fail to read the learning goal, the robot will read it for them.

### **Question 1:**

There are two students that are almost fighting. The child can get the teacher or to interfere themselves. Getting the teacher might result in being too late before the fighting begins.

Underlying values: Security/Tradition vs Universalism. It is safer to get the teacher, in addition, it is the norm to get the teacher. However, they might be too late and if they interfered, they possibly protected the welfare of both students.

### *English:*

Text before question: That's a great decision.

Clarification: Why did you choose this? <answer>

Implications & Consequences: What would the consequences of your action be? <answer>

Text after question: Thank you for explaining!

Tekst voor vraag: Dat is een goede beslissing.

Clarificatie: waarom heb je hiervoor gekozen? <antwoord>

Implicaties en consequenties: wat zouden de gevolgen van de keuze zijn? <antwoord>

Tekst na vraag: Bedankt voor de uitleg!

### **Question 2:**

The situation will be the same as before but now the teacher is further away and thus the child is sure that a fight has already started. This can impact the decision.

Underlying values: The same values as question 1 but we now ensure that they know it will be too late when the teacher is being called.

### *English:*

Text before question: You know that the teacher is at least 5 minutes away. If you get the teacher you would be definitely too late to prevent harm. Would this impact your decision? <answer>

Clarification: Why? <answer>

Implications & Consequences: What would be the consequences of doing that? <answer>

Text after question: Thank you!

*Dutch:*

Tekst voor vraag: Je weet dat de juf halen minimaal 5 minuten duurt . Als je de juf haalt, zou je zeker te laat zijn om een gevecht te voorkomen. Zou dit je beslissing veranderen? <antwoord>

Clarificatie: Waarom? <antwoord>

Implicaties en consequenties: wat zouden de gevolgen zijn als je dit doet? <antwoord>

Tekst na vraag: Bedankt!

### **Question 3:**

The situation changes slightly as the fighting students are now very young. The child can more easily interfere and stop the fight. Would the child change their decision now.

Underlying values: Power vs Tradition. Interfering is considered as Power as it gives the child control of the situation. Getting the teacher is Tradition as it is the norm taught in school.

*English:*

Text before question: In this case the students are 6 years old. You are your own age. If you were to interfere, you could definitely and easily stop the fight. Would you interfere or get the teacher? <answer>

Clarification: Why would you choose to tell or not to tell? <answer>

Implications & Consequences: What will happen if you tell it or not tell it? <answer>

Text after question: That's a great answer!

*Dutch:*

Tekst voor vraag: In dit geval zijn de vechtende kinderen 6 jaar oud. Je bent je eigen leeftijd. Als er tussen zou stappen, zou je het gevecht zeker en makkelijk kunnen stoppen. Zou je er tussen stappen of de juf halen? <antwoord>

Clarificatie: Waarom zou je dit wel of niet vertellen <antwoord>

Implicaties en consequenties: Wat zal er gebeuren als je het wel of niet vertelt? <antwoord>

Tekst na vraag: Dat is een geweldig antwoord!

### **Question 4:**

The question is whether the child will tell the teacher and get other kids in trouble or not.

Underlying values: Tradition vs Benevolence (vs Conformity). Telling the teacher will definitely impact the students negatively as they get in trouble for fighting. However, there is a social norm/tradition to tell the teacher. Some answers could also reflect conformity as not telling the teacher might disrupt and undermine smooth interaction and group functioning.

*English:*

Text before the question: In this final question, the fight is already over. You saw the kids fight. They are both a little bit hurt but nothing severe. Are you going to tell the teacher or not?

<answer>

Clarification: Why would you tell or not? <answer>

Implications & Consequences: What will be the implication of this choice? <answer>

Text after question: What a good answer!

*Dutch:*

Tekst voor vraag: In deze laatste vraag is het gevecht al voorbij. Je zag de kinderen vechten. Ze zijn allebei een beetje gewond, maar niet heel erg, gewoon een paar blauwe plekken. Ga je aan de juf vertellen dat ze aan het vechten waren of niet? <antwoord>

Clarificatie: Waarom zou je dit wel of niet vertellen<antwoord>

Implicaties en consequenties: Wat zal er gebeuren als je het wel of niet vertelt? <antwoord>

Tekst na vraag: Wat een goed antwoord!

**End:**

*English:*

Great, we are done with the third level, do you still remember the learning goal?

Yes: Wow, you are doing great, can you tell him to me?

Good answer: You remembered it very well!

Wrong answer: No problem, I remembered him as well, I will tell him to you

Some children have more difficulty understanding jokes than other children

No/Forgot: Okay, no worries, we are learning together and I remembered it as well, I will repeat it to you: Some children have more difficulty understanding jokes than other children.

Good answer: You remembered it very well!!

We are now done with the third couple of questions, you can click done in the game.

*Dutch:*

Super, we zijn nu klaar met het derde level, weet jij dit leerdoel nog?

Ja: Wow, wat goed van jou, kan jij hem aan mij vertellen?

Goed antwoord: Super onthouden

Verkeerd antwoord: Geen zorgen, ik heb hem ook onthouden, ik zal hem nog even aan je vertellen: Sommige kinderen snappen sommige grappen minder goed dan andere kinderen.

Nee/Vergeten: Oke, geen probleem, we leren samen en ik heb hem ook onthouden, dan herhaal ik hem wel even. Sommige kinderen snappen sommige grappen minder goed dan andere kinderen.

Goed antwoord: Jij hebt het leerdoel super goed onthouden!

We zijn nu klaar met de derde paar vragen, je mag op klaar klikken in het spel.

#### B.2.4 Situation 4

In this situation, the child gets grouped with a student that has reading and writing difficulties. They have a decision to take most of the work themselves to speed up the assignment or to split up the

work equally.

### Script

#### **During level:**

The robot will ask the child to read the learning goal. If the children fail to read the learning goal, the robot will read it for them.

#### **Question 1:**

The student makes the decision between splitting up the work equally or taking up most of the work

Underlying values: Self-direction vs Universalism. The child can choose to prioritize self-direction and take charge of the project or prioritize universalism and include their teammate and support them.

#### *English:*

Text before question: That's a good choice!

Clarification: Why did you choose this? <answer>

Implications & Consequences: What would be the implication of your choice? <answer>

Text after question: That makes sense!

#### *Dutch:*

Tekst voor vraag: Je hebt een goede keuze gemaakt.

Clarificatie: Waarom heb je dit gekozen? <antwoord>

Implicaties en consequenties: Wat zou het resultaat van je keuze zijn? <antwoord>

Tekst na vraag: Dat klinkt logisch!

#### **Question 2:**

In this scenario, the child works with Mike and two other students. Mike wants to work on the project in a creative way, your other group members want to stick to the normal method. The child can choose between the creative method, which conforms to Self-Direction or to stick to conforming with the group

Underlying values: Self-direction vs Conformity. Choosing to solve the problem in a creative manner is related to Self-direction, while conforming with the group is conformity.

#### *English:*

Text before question: You are now working in a group of 4 with Mike. Mike wants to do the assignment in a creative way while the other group members want to use the familiar method, do you choose the creative way or the way that the rest of the group suggested? <answer>

Clarification: Why? <answer>

Implications & Consequences: What would the other children and Mike think about your choice? <answer>

Text after question: Great answer!

*Dutch:*

Tekst voor vraag: Je zit nu in een groepje van totaal 4 kinderen met Mike er ook in. Mike komt met een creatieve oplossing voor de opdracht. De rest van het groepje wil het graag op een bekende manier oplossen. Kies jij voor de creatieve manier of de manier die jullie al vaker hebben gedaan? <antwoord>

Clarificatie: Waarom? <antwoord>

Implicaties en consequenties: Wat zouden de andere kinderen in je groepje vinden en wat zou Mike vinden van jouw keuze? <antwoord>

Tekst na vraag: Goed antwoord!

**Question 3:**

In this scenario, the child sees that Mike struggles with reading along. The child can ask the teacher to read more slow to accommodate Mike or say nothing to avoid potential backlash.

Underlying values: Benevolence and Security. Choosing to ask the teacher to read more slowly benefits the other student which is considered benevolence. Not facing the potential backlash is considered security.

*English:*

Text before question: You are now reading a book along with the class. You notice that Mike is struggling a bit with the pace of the class. Are you going to ask the teacher to read a bit slower, possibly facing some discomfort or are you not going to say anything? <answer>

Clarification: Why did you choose this? <answer>

Implications & Consequences: What would happen if you do this? <answer>

Text after question: I understand!

*Dutch:*

Tekst voor vraag: Je bent nu een boek met de klas aan het lezen. Je merkt dat Mike wat moeite heeft met hoe snel iedereen leest. Ga je aan de juf vragen om wat langzamer te lezen waardoor je je misschien wat ongemakkelijk voelt of ga je niets zeggen? <antwoord>

Clarificatie: waarom heb je dit gekozen? <antwoord>

Implicaties en consequenties: Wat zou er gebeuren als je dit doet? <antwoord>

Tekst na vraag: ik begrijp het!

**Question 4:**

The child is given the option between choosing something to improve themselves or something fun.

Underlying values: Stimulation vs Hedonism. If the person chooses fun it is considered Hedonism. Choosing something to improve themselves is considered Stimulation as it is a challenge.

*English:*

Text before question: For the group project, you can decide what you want to do. You can decide to do something that you need to learn more about because you think it's difficult, or you can also choose to do something fun. Do you want to improve yourself in something difficult or do something fun? <answer>

Clarification: Why did you choose this option? <answer>

Implications & Consequences: What would be the implication of this choice? <answer>

Text after question: Thanks!

*Dutch:*

Tekst voor vraag: Voor het groepsproject mag je kiezen wat je wilt doen. Je kan kiezen om iets te doen waar je meer over moet leren omdat je het moeilijk vindt, of je kan iets doen wat je leuk vindt. Kies je voor jezelf verbeteren in iets moeilijks of iets leuks doen? <antwoord>

Clarificatie: Waarom heb je hiervoor gekozen? <antwoord>

Implicaties en consequenties: wat zou het gevolg van deze keuze zijn? <antwoord>

Tekst na vraag: Bedankt!

**End:**

*English:*

Great, we are now done with the last level, do you still remember the learning goal?

Yes: Wow, you are doing great, can you tell it to me?

Good answer: Wow, you do not even need me!

Wrong answer: Okay, you remembered it quite good already, I remembered this:

Some children have to put more effort in reading and writing without mistakes than others.

No/forgot: Okay, no problem, I remembered this: Some children have to put more effort in reading and writing without mistakes than others. .

Good answer: Wow, you do not even need me!

We are done with the last few questions, can now click Done in the game.

*Dutch:*

Super, we zijn nu klaar met het laatste level, weet jij het leerdoel nog?

Ja: Wow, wat goed van jou, kan jij hem voor mij opnoemen?

Goed antwoord: Wow, jij hebt mij niet eens nodig

Verkeerd antwoord: Oke goed onthouden, ik heb dit onthouden: Sommige kinderen moeten meer werk steken in foutloos lezen en schrijven dan andere kinderen.

Nee/Vergeten: Oke, geen probleem, ik heb dit onthouden: Sommige kinderen moeten meer werk steken in foutloos lezen en schrijven dan andere kinderen.

Goed antwoord: Wow, jij hebt mij niet eens nodig!

We zijn nu klaar met de laatste paar vragen, je mag op klaar klikken in het spel.



## **C Information form and consent form**

### **C.1 Information form parents**

# Informatiebrief ouders/voogd – Onderzoek: Bewustzijn over neurodiversiteit creëren door middel van het spelen van een educatief spel met een robot.

## Wie ben ik?

Ik ben Laura Ottevanger, een masterstudent van de TU Delft. Ik studeer Informatica. Voor het afronden van een master studie dient er een onderzoek gedaan te worden. Ik ben de zus van juf Sandra Ottevanger.

## Wat onderzoek ik?

Op meerdere plekken wordt momenteel onderzocht hoe robots kunnen worden ingezet voor het onderwijs. Ik richt me op robots die kinderen helpen om individueel te werken aan een leeropdracht op de computer. De kinderen leren specifiek over neurodiversiteit en over hun waardes (wat ze belangrijk vinden, bijvoorbeeld bij het maken van contacten in de klas).

Hiervoor hebben we de hulp van uw kind nodig. Alle leerlingen van de klas worden gevraagd om mee te doen aan dit onderzoek. Ze zullen uitgenodigd worden om onder schooltijd een half uurtje met de robot een spel te spelen. Aan het einde van de sessie meten wij of de gesprekken met de robot goed hebben gewerkt. Met deze uitkomsten willen wij de robot en diens gesprekstechnieken verbeteren.

## Wat moet uw kind doen en wat doen wij?

Wij willen weten of een robot goed kan helpen bij het behalen van leerdoelen over neurodiversiteit en hoe hiermee inzicht gegeven kan worden in de waardes van kinderen.

Aan de hand van een sessie van ongeveer 30 minuten met uw kind en de robot verzamelen we die gegevens. Er wordt een geluidsopname gemaakt van de sessie. Deze geluidsopname wordt uitgetypt, geanonimiseerd en daarna verwijderd.

De sessie zal bestaan uit het spelen van een spel. Uw kind zal een opdracht krijgen, een voorbeeld hiervan is het organiseren van een feestje. Vervolgens krijgt uw kind de relevante leerscene in het spel aangeboden. In dit voorbeeld moet er een DJ worden gevonden voor het feestje. Er zijn twee mogelijkheden: een kind die een hele goede DJ is maar snel overprikkeld raakt in drukke situaties en een kind die heel erg graag DJ wilt zijn. In deze situatie moet er dus tussen deze twee kinderen een keuze worden gemaakt. Aan de hand van deze keuze zal de robot een vraag stellen naar het waarom van deze keuze of naar de mogelijke consequenties van de keuze. Voor het onderzoek test ik hiermee welk soort vraag beter is om iemand bewust te maken van de waardes achter een keuze. Aan het einde van de sessie krijgt uw kind een aantal vragen over de stof, de keuzes en de robot.

## Hoe doet uw kind mee?

U en uw kind mogen zelf weten of uw kind meedoet. Als uw kind mee wilt doen, vragen wij toestemming van uw kind en van u. Er zitten geen consequenties aan of uw kind wel of niet mee doet.

## Welke gegevens van uw kind worden opgeslagen?

Van uw kind wordt alleen opgeslagen in welke groep uw kind zit. Verder wordt er een geluidsopname gemaakt van de sessie met de robot, die wordt uitgetypt en geanonimiseerd en daarna verwijderd. De resultaten en gegevens die worden opgeslagen zullen verder niet gedeeld worden met de docenten van uw school.

## Wat gebeurt er met de gegevens als het onderzoek is afgelopen?

De audiotranscriptie van uw kind wordt op het beschermde netwerk van de TU Delft bewaard.

U en uw kind kunnen voor het einde van het onderzoek (1 juni) te allen tijde uw toestemming intrekken, dan worden alle verzamelde gegevens van uw kind verwijderd. Na het einde van het onderzoek zullen de gegevens volledig zijn geanonimiseerd en is het niet meer mogelijk om te onderscheiden welke data van uw kind is.

### Waarom zou uw kind meedoen?

Uw kind krijgt de kans om met een NAO-robot te praten.

Uw kind leert over neurodiversiteit.

Uw kind draagt bij aan een onderzoek om robots in de klas te verbeteren.

### Waarom zou uw kind niet meedoen?

Uw kind mist ongeveer 30 minuten van de normale lestijd.

### Hoe beschermen we de privacy van uw kind?

Wij slaan alleen de groep van uw kind op en de uitgetypte versie van de geluidsopname, verder worden persoonlijke gegevens niet opgeslagen.

Als uw kind tijdens de sessie een naam noemt, dan wordt deze vervangen door een nep naam.

De gegevens blijven geheim voor anderen. Alleen de onderzoeker heeft toegang.

Alle gegevens van uw kind krijgen een code die niets betekent (bijvoorbeeld C1). Zo kan niemand achterhalen wie welke code is. De data zullen na het onderzoek volledig worden geanonimiseerd, hierna is het onmogelijk om te achterhalen welke data van welk kind zijn.

De data zoals de uitgetypte, geanonimiseerde versie van de geluidsopname kunnen nog worden gebruikt bij de research groep maar zal verder niet gebruikt worden buiten de research groep om. Deze uitgetypte versie van de geluidsopname wordt verder niet gepubliceerd als dataset.

De gegevens worden vernietigd na 10 jaar.

### Verdere vragen?

Mocht u of uw kind verdere vragen hebben dan kunt u mailen met Laura Ottevanger:

[l.c.ottevanger@student.tudelft.nl](mailto:l.c.ottevanger@student.tudelft.nl)

Delft University of Technology  
**Toestemmingsverklaring**  
*Informed consent form*

Ondergetekende partijen geven toestemming voor deelname van zijn/haar kind aan het onderzoek getiteld *Bewustzijn over neurodiversiteit creëren door middel van het spelen van een educatief spel met een robot*.  
*The undersigned parties consent to his/her child's participation in the study entitled Raising awareness about neurodiversity by playing an educational game with a robot*

.....

Naam kind .....

Ik bevestig dat ik de informatie over bovengenoemd onderzoek heb gelezen. Ik begrijp de informatie.  
*I confirm that I have read the information about the study mentioned above. I understand the information.*

De bedoelingen van het onderzoek en de daarbij gevolgde aanpak zijn tot mijn tevredenheid uitgelegd.  
*The intentions of the study and the approach followed have been explained to my satisfaction.*

Ik heb de gelegenheid gehad om aanvullende vragen te stellen en deze vragen zijn naar tevredenheid beantwoord.  
*I have had the opportunity to ask additional questions, which have been answered satisfactorily.*

Ik heb voldoende tijd gehad om over deelname van mijn kind na te denken.  
*I have had sufficient time to think about the participation of my child.*

Ik weet dat de deelname van mijn kind aan het onderzoek geheel vrijwillig is en dat ik de deelname op ieder moment kan beëindigen voor het einde van de studie (1 juni) zonder dat ik daarvoor een reden hoeft op te geven.  
*I know thamy child's participation in the study is completely voluntary and that I can withdraw my consent until after the end of the project (1<sup>st</sup> of June) without providing a reason.*

Ik geef toestemming voor de verwerking van de persoonsgegevens van mijn kind voor de doelen zoals beschreven in de informatie.  
*I give permission to process the personal data of my child for the purposes described in the information.*

Ik geef toestemming dat quotes of opmerkingen van mijn kind kunnen anoniem in het verslag worden gezet.  
*I give permission to use quotes or responses from my child to be used anonymously in the report.*

Ik geef toestemming de onderzoeksgegevens van mijn kind te hergebruiken voor toekomstig onderzoek op het beschreven onderzoeksgebied op voorwaarde dat deze zo gecodeerd zijn, dat ze niet meer naar mijn kind terug te leiden zijn.  
*I give permission to reuse the research data of my child for future research in the research area described, provided that these are coded in such a way that they can no longer be traced back to my child.*

Ik geef toestemming voor het bewaren van de gegevens van mijn kind en dat bevoegde leden van het onderzoeksteam en bevoegde inspecteurs hier inzage in hebben.  
*I give permission for my child's data to be stored and for authorized members of the study team and authorized inspectors to have access to the data.*

Het toestemmingsformulier moet worden getekend door een gezaghebbende ouder of door de voogd  
*The informed consent form must be signed by an authoritative parent or by the guardian.*

Naam, handtekening en datum ondertekening gezaghebbende ouder/voogd  
*Name, signature and date of signing of authoritative parent/guardian*

.....

Voor de onderzoeker:

Ik heb verzekerd dat ik de ouders / de voogd goed geïnformeerd heb over het onderzoek waaraan het kind gaat deelnemen. Ik heb de ouders/ de voogd en kinderen laten weten waar ze vrijwillig toestemming voor geven.  
*I have ensured that I have properly informed the parents/guardian about the study in which the child will participate. I informed the children and parents/guardians to what they are freely consenting.*

Naam, handtekening en datum ondertekening proefleider:  
*Name, signature and date of signing study leader*

.....

## C.2 Information form children

## Informatie over het onderzoek: Bewustzijn over neurodiversiteit creëren door middel van het spelen van een educatief spel met een robot

### Waar gaat het over?

We willen je vragen of je mee wilt doen aan een onderzoek naar robots in de klas. Dit onderzoek wordt gedaan door Laura Ottevanger van de TU Delft. We willen kijken of robots kunnen helpen om beter te leren. Het onderzoek gaat over neurodiversiteit, voorbeelden daarvan zijn ADHD en Dyslexie. Als je wilt weten wat dit is, dan kan je dit aan mij of aan de juf vragen.

Je mag zelf beslissen of je meedoet.

### Wat ga jij doen?

Je gaat op school samen met de robot leren. Jullie gaan samen een spelletje spelen. De robot zal je af en toe vragen stellen. Daarna vul je een vragenlijst in.

### Wanneer en hoe lang?

Je gaat 1 keer met de robot spelen en praten. Dit duurt ongeveer 30 minuten. Het is tijdens schooltijd.

### Voordelen en nadelen?

- + Je gaat praten met een coole robot.
- + Je helpt mee aan het beter maken van de robot. Zo kan hij nog beter worden in het helpen met leren!
- Je bent dertig minuten weg uit de klas. Misschien mis je een stukje van de les.

### Belangrijk om te weten

- Je hoeft niet mee te doen als je niet wilt
- Je mag altijd stoppen zonder dat je hoeft te vertellen waarom
- Je mag altijd vragen stellen
- Of je mee doet of niet heeft geen invloed op je rapport/cijfers
- Je krijgt geen cijfer en wordt niet beoordeeld
- De juf(fen)/meester(s) krijgen niet te horen over wat jij hebt gezegd/gedaan met de robot

### Als je vragen hebt

Vragen kun je met je ouders bespreken. Je kunt ze samen aan Laura Ottevanger stellen.  
Je kunt Laura mailen op: [l.c.ottevanger@student.tudelft.nl](mailto:l.c.ottevanger@student.tudelft.nl)

Als je vragen hebt dan mag je ze hier schrijven:

### Toestemming

- Ik heb alles gelezen
- Ik snap alles
- Ik wil meedoen

Schrijf hier je naam op: \_\_\_\_\_



### C.3 Benefits of gaming literature review

In the present day and age, a lot of how our time is spent is dominated by looking at a screen. The amount of people using the internet is 5.07 billion people around the world [60]. The number of video game users reached 3 billion for the first time in 2020. In the Netherlands, 74.3% of all internet users are reported to play video games according to a survey done by in the third quarter of 2021. Out of all of the survey respondents between the ages of 16 and 64 years old, the youngest age range (16-24) are reported as the highest percentage of video-gamers as 86,4% of the females in the age range and 91.1% of the males have reported to play video games. A survey done by the Trimbos Instituut reports that 89% of the students attending primary school in the Netherlands, occasionally play a video game. To summarize all of these numbers, a lot of people, especially children and young adults play video games. These numbers show that the outreach to children can be massive when video games are used as a medium.

A lot of research done on the effects of video games are often focused on the negative impact. However, there can also be great benefits to gaming. Granic et al report a lot of the benefits to gaming [61]. First of all they reported cognitive benefits of gaming. For example, people who play shooter games often show positive effects in tasks where top-down attention or the processes that control and regulate attentional allocation and resource management are required. In addition, enhancements in task-switching, multi-tasking, and visual short-term memory tasks have also been shown [62]. A meta-analysis showed that improvements of spatial skills that are derived from playing shooter games are comparable to high school and university-level courses aimed at enhancing these skills. The meta-analysis showed that spatial skills can be trained with video games in a short amount of time. In only 10 hours, the participants playing the shooter games showed a significant improvement in their spatial thinking skills [63].

Video gaming can also be associated with enhanced creativity. Jackson et al. showed a relationship between playing video games and enhanced creativity in children. Other technology did not show the same relationships. However, it is not clear if this relationship is causal and in what direction it would be causal [64].

Granic et al. also reported the motivational benefits of gaming. They propose that video games are a great training ground for acquiring an incremental theory of intelligence as they provide concrete and immediate feedback on the specific efforts of the players. The flow of a video game can also create motivation. When boredom or anxiety increases, the player is motivated to strive for the flow state in order to feel enjoyment again [57].

The next benefits reported by Granic et al. were emotional benefits. They mention several studies with a causal relation between playing video gaming and increases in positive emotion and improved mood. [65][66]. The previously mentioned flow experience can also be linked to positive emotions; higher commitment, achievement, self-esteem and less anxiety [67].

Finally, Granic et al. reported Social benefits. They propose that gamers are quickly learning social skills and prosocial behaviour that are useful outside of the games. Social skills are also improved due to civic engagement; organizing groups and leading people in social causes.

A game is free to show metaphors. Rieber and Noah did a study where the participants played a game and were given different levels of metaphors and information. They showed that using the combination of visual metaphors and a game, the participants were able to transfer their understanding to another game context [68]. The metaphor by itself, without the game, had no effect on explicit learning.

In conclusion, video games are a widely used form of entertainment, with billions of users worldwide. While there is much research focusing on the negative aspects, it is important to note the potential benefits of gaming. There are many cognitive advantages like improved attention, task-switching and spatial skills. Gaming also promotes creativity, and motivation through incremental learning

and positive emotions. There are also social benefits due to the development of social skills and civic engagement. In addition, videogames are free to use metaphors to enhance understanding and learning. These benefits can be used to inform the development of educational games.

## D Game images



Figure 27: The character selection screen. The player can input their own name and make themselves.



Figure 28: The first level. The learning goal is at the bottom of the screen. The action that the player should do is at the top of the screen.





Figure 29: The first level. interaction with the teacher. The teacher gives the next task to the player



Figure 30: The first level. the second learning goal is at the bottom of the screen. The new task is at the top of the screen.



Figure 31: The first level. The conversation with one of the two peers.



Figure 32: The first level. The decision that the player makes based on previous conversations and their values.



Figure 33: The second level. Introduction to the problem by a NPC.



Figure 34: The second level. The moral decision that the player makes based on their values.



Figure 35: After every level. The game leaves space for the robot to ask questions.