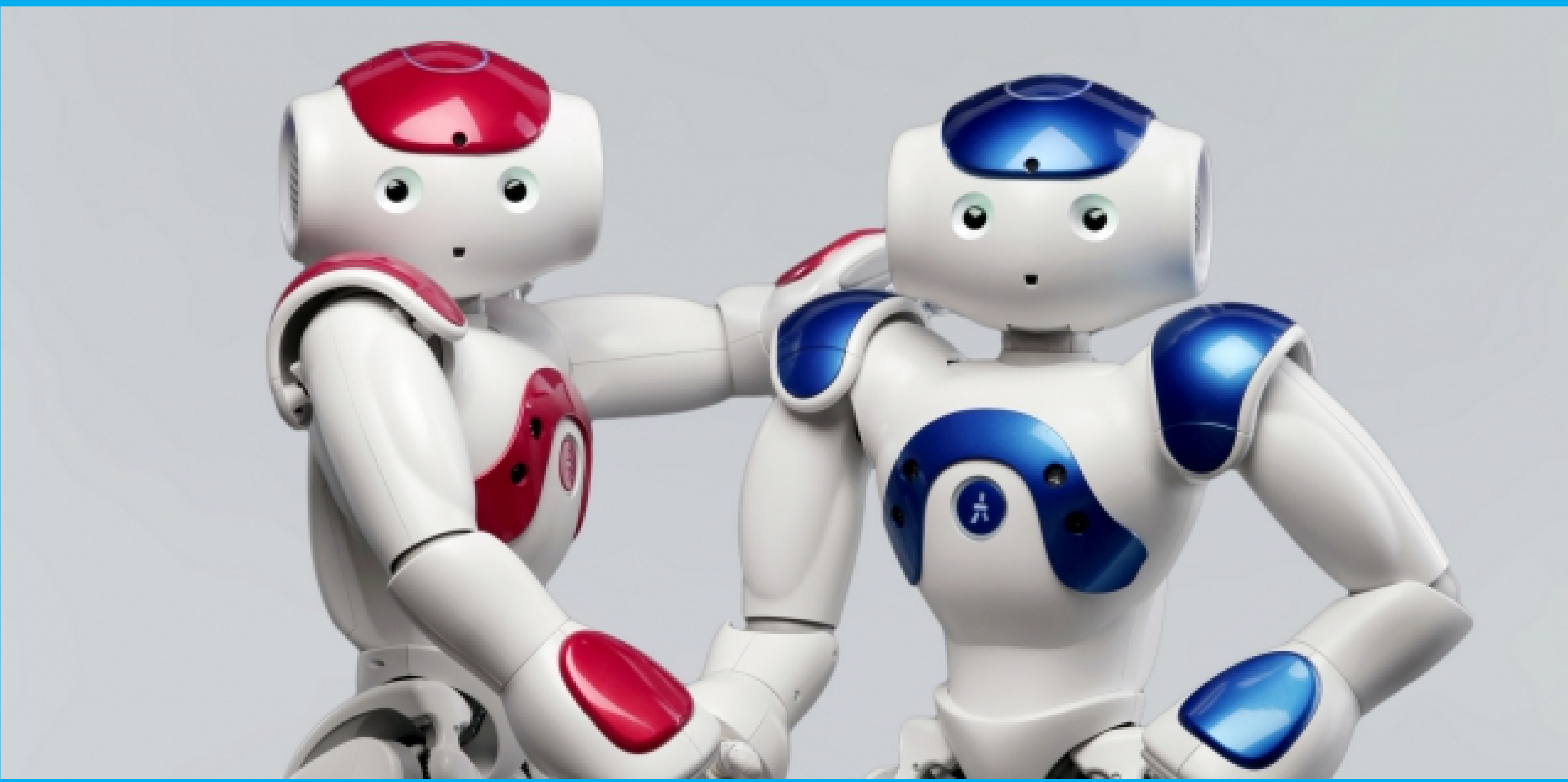
Adaptive feedback by a humanoid robot tutoring math

S.P.T.M. Liebens



TUDelft

Adaptive feedback by a humanoid robot tutoring math

by

S.P.T.M. Liebens

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday February 26, 2019 at 15:00 PM.

Student number:	4207750	
Project duration:	September 1, 2017 – February 26, 2019	
Thesis committee:	Dr. K. V. Hindriks,	TU Delft, supervisor
	Drs. M. A. F. M. Jacobs,	TU Delft
	Prof. Dr. M. A. Neerincx	TU Delft

TU Delft

Abstract

Schools and other educational instances (such as: elementary, primary and secondary schools, schools for applied sciences and universities) are making use of new innovations and technologies to help teachers and students. After Smart Boards, computers, and tablets, a new field of research arises: can robots in any way help in teaching? During this research an answer to the question ”What is the effect of elaborated feedback versus minimal feedback given by a humanoid robot (NAO), on a primary school student solving basic math problems?”. Also, two hypothesis are tested: ”Feedback provided by NAO will improve the student math test results.” and ”Feedback improved the affection of the student towards NAO.”. We find that lacing and split and add are the most used calculation strategies, and NAO also aims to identify commonly made mistakes such as switching units in bigger number or confusion about symbols or strategies. An important element of this research is feedback, which concerns information about how we perform in efforts to reach goals. It requires information about the goal, the followed track and the next steps. Feedback seems to be most effective when it is self-regulated and intended for process. Children need tangible and timely feedback, not too early (to prevent overwhelming the student) and not too late. Giving feedback on the result as well as on the track, challenging students and defining a clear road to success are crucial in helping the student succeed. In this research, NAO will provide this feedback and thereby should help the student succeed in solving math problems.

Two scenarios are designed: an introductory class, to let students touch, feel and experiment with NAO, and a math tutoring session, which will be used during the experiment. We conduct an experiment, testing a feedback group versus a control group, where the feedback group is able to get feedback twice (elaborated feedback), and the control group only hears whether the answer given was wrong or correct. The elaborated feedback consists of detecting a mistake made by the student, and adapt the feedback according to the given answer. After a second try and incorrect answer, the process to the correct answer is explained to the student. Students had to fill in a PANAS form in order to measure affection towards NAO during the sessions. No significant results were found in both the math test and the PANAS scores. We did however find that students learned to work with NAO very quickly, interaction was robust, although speech recognition needs some improvement. Interesting to see was that we found contradictions when we compared the qualitative analysis with the quantitative analysis. Statistical analysis showed that there was no significant growth in the skills of a student, but all students mentioned that they felt that they did learn from the session with NAO.

Future work could involve conducting the experiment for a longer period of time, to see a better lasting effect of the feedback given. Other work could be done in distinguishing groups according to their level of math, to see if students with a low grade in math have a higher profit from NAO’s influence. Also, combining mistakes could lead to more mistakes that can be detected, and thus feedback could be more efficient.

Preface

Before you start reading this research, there are a few things you must know. Firstly, this is a research involving a robot, a humanoid robot: NAO. Although it has no human sense, or human capabilities at all, in this paper some references to human capabilities might be found in combination with the robot. This is done, because we humans like to give human capabilities to objects. So in this paper, NAO understands, hears, and listens. At least, I like to believe it does. Secondly, this thesis was a fun, terrifying and amazing process to be involved in. I learned a lot about myself, my way of working, conflicting priorities, places to study and concentration spans.

It all started with ”I want to do something with robots”. That changed quickly to a choice between robots and elderly people or robots and children. For me, the choice was easy: as a volleyball trainer, I love working with kids, so children it was! Next off: a subject. Robots in primary schools became our project name. Where my first idea was to create an adaptive calculation robot, it soon became clear that feedback was a more interesting field of research. At IPON, an innovation exposition for educational instances, I found that feedback is used too little in current applications, but thought that this could increase the performance of the students. So my road became clear: literature about feedback, the dutch teaching system, and robots in education. A stressful summer of programming NAO, to actually provide the feedback, and 4 great weeks of working with NAO and the students.

My road to success has been a rollercoaster, with happy moments, and sad moments. Moments of full panic, and moments of pure joy. Many times I thought of the quote from Pierre de Fermat, in Fermat’s Last Theorem: ”The margin of this page is a bit too small to write down the proof”, had a laugh, and continued writing.

During my road to success, I had a lot of help. I want to thank Mariëlle van Es, for sharing her knowledge as a teacher at primary schools. A special thanks to Katelijne and Tineke, for using their fifth grade students in my research. I also want to thank Koen and Martin, my two project supervisors, who had to deal with me, my spam and my work for over a year. And I would like to thank my friends and family for supporting me throughout this entire process. It was a pleasant ride, I hope you enjoy reading this thesis as much as I enjoyed writing it.

S.P.T.M. Liebens
Delft, February 2019

Contents

1 Introduction

Schools and other educational instances (such as: elementary, primary and secondary schools, schools for applied sciences and universities) are making use of new innovations and technologies to help teachers and students. After Smart Boards, computers, and tablets, a new field of research arises: can robots in any way help in teaching? NAO robot is a humanoid robot, that has already been put to use in educational instances as a help to students. A picture of NAO can be seen in figure 1.1. More details about NAO can be found in section 5.3.

Figure 1.1: NAO robot.

As the interest of teachers and educational instances in using innovations to help in teaching grows, questions arise in the use of NAO robots in these instances. In this study, the use of NAO as a teaching aid in math classes will be explored.

Before the age of 4, children start learning simple calculations by playing games or asking parents how many spoons have to be on the table. This leads to incidental knowledge: knowledge that is gained by the child without direct teaching. This incidental knowledge will grow especially before children start at third grade (age 6 and 7) in primary school, after which math education by teachers will start [20].

In the Netherlands, primary and secondary schools make use of SLO standards (`http://www.slo.nl/`). SLO is the Netherlands Institute for Curriculum Development. Its tasks are to design and validate national curricular frameworks (core objectives, attainment levels, examination programmes, curricular strands). In this research with NAO, focus will be put on core objective 27: The student learns to do basic operations on natural numbers up and until 100, quickly and without the use of paper or other tools, of which addition and subtraction until 20, and basic multiplications are known by heart [5].

In order for children to learn, motivation and feedback are of the essence. A motivated child has a better focus and will be able to remember more of the taught matter. Verschuren has done a lot of research in Dutch primary schools about motivation and feedback. Motivating children the correct way, and giving timed feedback is an important part of teaching [40].

Feedback can be found in many forms. One can praise the student, give advice or evaluate the current situation. According to Hattie and Timperley [12], feedback should always contain three questions: Where am I going? How am I going? And where to next? These questions should be intervened with four levels of feedback, in order to give correct feedback on the task. It can either be about the task, about the process, about self-regulated feedback or about praising a person. Besides these questions and levels, timing is of the essence when giving feedback. Immediate feedback gives better results than waiting for a long time, but still immediate feedback can be overwhelming. Feedback should be timed correctly, according to the task that has been done [33, 43].

Current math applications, such as De Rekentuin (`https://www.rekentuin.nl/`) and Snappet (`https://nl.snappet.org/`), work on tablets or computers. These applications work in a gamified way. They give coins or presents when students give a correct answer and simply continue when students give a wrong answer. These applications provide teachers with learning analytics such as which exercises were answered correctly, and after how many tries, and which exercises were answered wrongly, but also the amount of coins or presents a student has earned. This amount represents the level of the student by a method developed by the app developers.

These apps, however, show only learning analytics to a teacher, who has to take actions accordingly him- or herself. Although these apps already show a teacher where to focus at for which student, the teacher still has to take time to provide information to the student about the mistakes (s)he has made, and where in the calculation the mistake is made, so that in future practices it will not happen again.

What if we were able to use NAO in such a way, that not only it provides the student with exercises, but also provides the student with feedback about its mistakes? Could the student learn from feedback provided by NAO?

In order to achieve that, NAO must be able to recognise wrong answers. To do so, a list of commonly made mistakes will have to be made, which NAO can use. An experiment will be conducted, that will test a testing group versus a control group of students. NAO will provide some exercises, after which the testing group will get elaborated feedback when an answer has been given. The second session, the groups will get excercises of categories that went wrong the first time, to see whether NAO's feedback has had effect.

In future, robots might even be able to really assist teachers in what a student needs. Really adapt to a student's level of knowledge, help accordingly, and adjust feedback per person. Teachers and robots could work together as a team, where all student data is gathered in a cloud, to be accessed for everyone who works with the student. This could reduce the stress and workload of a teacher, so that he or she can really invest his or her time in what is important: the student's goal, growth and learning capabilities.

The main question in this research is: **What is the effect of elaborated feedback versus minimal feedback, given by a humanoid robot, on a primary school student solving basic math problems?** In order to create a robot which can give appropriate feedback to the student, some building blocks are needed. To do research to and development of a robot that would be able to help students with math, and provide feedback, we can ask three subquestions according to specific subjects of the process: Research, Design and Effects. The first question will mainly focus on the theory and literature. It will also address problems or questions that apply to NAO before we make design choices. The second question addresses some design choices that have to be made for NAO to function correctly. The third question is about the effects of NAO's feedback on the students and teachers. To answer that question, we can create a survey with children for after the sessions, and talk to teachers about their experiences.

1. In order to provide feedback, we need to know what mistakes are made by students, what calculating strategies are taught, and when and how to provide feedback. We need to find an answer to the research question: *How can a humanoid robot provide feedback after a mistake was made by a primary school student in a basic calculation exercise with natural numbers up until 100?* We will find answers to that question by doing a literature research, searching for the subjects feedback, calculating strategies, and thinking processes of students.

2. A robot can be very impressive for students, and we need the robot to be educational and helpful. We need to know how NAO should act and behave in order to succeed in its task. *How will the interaction mechanism of NAO be shaped, in order to communicate with the student in the tutoring session?* To find an answer to this question, we will search in literature about how Human-Robot-Interaction is mostly done in combination with students, how rewards and human-like behavior can be adapted into NAO. Also, we will need to know how NAO can fulfil the role as a tutor.

3. When we know what effects the robot and feedback have on the students, it could be developed further in order to be of added value for educational instances. In this research, an answer to the question below is found: *How did feedback and the robot influence the performance, affection and interaction of the student in practising math?* By conducting an experiment involving NAO in a primary school, we try to find an answer to the question above.

During the analysis of the gathered data, we will test two hypotheses:

- Feedback provided by NAO will improve the student's math test results.

- Feedback increased the affection of the student towards NAO.

This thesis report will guide you through the process of letting NAO provide elaborated feedback to primary school students who practise math. First of all, in chapter 2, some background information will be provided about feedback, teaching in primary schools, calculation methods and robots in education. Chapter 3 will mention the design of NAO, design of the feedback and other design choices that have been made. Chapter 4 will guide you through the process of development. What algorithms and development tools have been used to provide NAO with instructions to succeed? Chapter 5 contains the method of the experiment, in which the experiment will be explained, together with the research tools and participants. The results of the conducted experiment will be analysed in chapter 6, and lead to a discussion and conclusion in chapters 7 and 8 respectively.

2
Background

NAO might be able to help students in math by providing feedback to an answer. In order to achieve that, some background knowledge is needed. How is learning and teaching in the Netherlands organised? What calculation methods are taught to students? What are current teaching assistants and how are robots put to practise in education until now? Theoretical knowledge about feedback and motivation is also needed. This chapter will mention all theoretical background needed in order to answer our research questions.

2.1. Teaching Mathematics to Children

Before the age of 4, children start learning simple calculations by playing games or asking parents how many spoons have to be on the table. This leads to incidental knowledge: knowledge that is gained by the child without teaching. This incidental knowledge will grow especially before children start at third grade (age 6 and 7) in primary school, after which math education by teachers will start [20].

2.1.1. The Dutch System

Current primary and secondary schools in the Netherlands make use of SLO terms. (`http://www.slo.nl/`) SLO is the Netherlands Institute for Curriculum Development. Its tasks are to design and validate national curricular frameworks (core objectives, attainment levels, examination programmes, curricular strands). They create a curriculum containing skills and methods that students should know at the end of a learning year (attainment levels) These curricula contain core objectives (Dutch: kerndoelen), and are divided into several groups that belong to the same aspect. Math for example, has an aspect ”Numbers and operations”, which contains all core objectives that have an impact on numbers and operations on numbers. In this research with NAO, focus will be put on core objective 27: The student learns to do basic operations on natural numbers up and until 100, quickly and without the use of paper or other tools, of which addition and subtraction until 20, and basic multiplications are known by heart [5]. To be more precise, this research with NAO focusses on: addition and subtraction of numbers up and until 100. This core objective is put to use mostly in 4th grade (age 7 and 8). So in order to see if students memorise (known by heart), understand or apply the taught matter, and to see if the NAO robot can help the students in learning it, our research will focus on 5th grade (age 8 and 9). According to Soest, students in fifth grade can be expected to quickly add and subtract numbers up and until 100 [39].

Each core objective has a learning curve, divided over 4 sections. The first section is for grade 1 and 2 (age 4-6), the second section for grade 3 and 4 (age 6 - 8), the third section for grade 5 and 6 (age 8 - 10) and the last section for grade 7 and 8 (age 10 - 12). Concerning core objective 27, in grade 1 and 2 the students start with numbers as an amount, in for example passengers on a train, amount of cookies in a jar or parking cars. They also use their fingers to add and subtract numbers. In third and fourth grade, students start using calculation methods as can be found in section 2.1.2. In fifth and sixth grade, these methods are repeated, and teachers start using bigger numbers, until in grade 7 and 8, the students should easily be able to add and subtract any numbers by heart.

2.1.2. Calculation Methods

SLO makes a distinction between ”simple strategies” and ”calculation strategies”. Simple strategies are used for basic calculations with numbers smaller than 10. Calculation strategies are being used in more extended calculations with numbers bigger than 10.

Simple strategies:

- Exchanging: 3 + 6 = 6 + 3
- Near Double: 6 + 5 = 5 + 5 + 1
- Transform: 4 + 6 = 5 + 5
- Compensate: 5 + 8 = 5 + 10 - 2
- Calculation via 10s: 6 + 8 = 6 + 4 + 4
- Inversion: 12 - 6 = 6 because 6 + 6 = 12

Calculation strategies:

- Lacing (*Dutch: Rijgaanpak*): 34 + 27 is done by first adding 20 to 34 = 54, then adding 7 to 54 = 61.
- Split and add (*Dutch: Splitsaanpak*): 34 + 27 is done by first adding 20 to 30 = 50, followed by 4 + 7 = 11. In the end add all together: 50 + 11 = 61.
- Compensating: 67 - 19 is done by exchanging the 19 for 20. 67 - 20 = 47, but we subtracted a 1 too much, so 47 + 1 = 48.
- Complementing: 50 - 48 is done by counting how much of a difference there is between 48 and 50.

According to research done by CITO, a knowledge organisation specialised in creating tests and exams for educational instances, the mostly used and taught strategy in primary schools is lacing, after which splitting is used the most secondly [19]. Wolsink and Os however, state that the used method is dependent on the method used by the school, and what method the teacher prefers. They also seem to see that splitting is more oftenly used rather than lacing[26, 44]. Although current literature does not answer the question about which strategy is mostly used, it can be stated that lacing and splitting are used mostly. During the implementation of NAO, focus will therefore be put on these two methods.

2.1.3. Commonly Made Mistakes

Mathematical knowledge is being developed according to four main procedures:

1. Conceptualisation, learning to give meaning to knowledge and skills.
2. Development of solutions.
3. Learning to quickly solve exercises (practice, automate, memorise).
4. Being able to apply knowledge and skills in a flexible way.

Some problems that may arise are for example: Does the student understand the context in which the exercise might apply? Or vice versa, does the student understand the symbols of the exercise, and is (s)he able to apply this to a context? Other problems that can arise are not necessarily problems of mathematical knowledge. When a student has a lack of knowledge in language or cannot read well, is not able to visualise the context, or cannot understand or memorise solutions to problems in the same category, the student will most likely also have problems with math.

Children who do not have complete knowledge of structures of tens and units might switch units in bigger numbers. For example, 23 becomes 32.

In subtraction, using the split and add method, most mistakes are made in the last step, because it is being confused with the addition step. When we try to solve 42 + 36, we first solve 40 + 30, which is 70, after which we solve 6 + 2, which is 8. Then we add 70 to 8 and we conclude the answer is 78. When we try to solve 42 - 36, we first do 40 - 30, which is 10, then the mistake is made in subtracting 2 from 6, instead of subtracting 6 from 2. 6 - 2 is 4, which leads to an answer of 10 + 4 = 14. Confusion whether to subtract or add, or which number to subtract from which number in the last step is a commonly made mistake in the split and add method. Less mistakes are being made using lacing, because the first number is not split [21]. In order to give correct feedback, NAO will need to recognise this mistake.

2.1.4. Solutions

According to research by Soest et. al. [39], several solutions to calculation problems and dyscalculia are relevant to students of this age. Amongst other things, he mentions:

- Let the student practice a lot with the exercises that went wrong, the more practice, the better.
- Give immediate feedback.
- Not only mention the result, explain out loud how the solution can be found.
- Cheer and give positive feedback.

Children in the age of 6 to 9 who have troubles with math, will need an environment without stress and distraction, but with full motivation and attention. NAO can provide this situation for children who are making a lot of mistakes in math.

2.2. Current Teaching Assistants

Current teaching aids such as Snappet or De Rekentuin use tablets or computers as a source of input for students. This section will show what current applications like Snappet and De Rekentuin actually do, and where improvements can be made.

2.2.1. Snappet

Adaptive educational technology tool Snappet combines extracted and embedded learning analytics in classrooms daily. The tool provides students with math exercises in a gamified way. Snappet gives coins or presents when students give a correct answer and simply continues when students give a wrong answer. Snappet applications provide teachers with analytics such as which exercises were answered correctly, and after how many tries, which exercises were answered wrongly, but also the amount of coins or presents a student has earned. This amount represents the level of the student by a method developed by the app developers. While students make exercises on the tablet, this technology displays real-time data of learner performance in a teacher dashboard (extracted analytics). At the same time, learner performance is used to adaptively adjust exercises to students' progress (embedded analytics). Teachers use this information as input to adjust instructions, modelling and feedback for the student, as well as selection of classroom wide practices [22, 23].

2.2.2. De Rekentuin

Like Snappet, De Rekentuin provides a gamified way of learning math for students. Students have to work in an online garden and can grow all sorts of plants and flowers by doing calculation games. The more a student plays, the more items it gathers to maintain its flowers. The better the student performs, the better the quality of the items. Responses of the students are, as in Snappet, automatically analysed and sent to the teacher. The teacher receives learning curves, number of mistakes made, and other comparisons and analytics [36, 37].

NAO can improve the situation of needing a teacher to act towards the learning analytics that are provided by De Rekentuin as well as Snappet. By providing feedback to the answer that a student has given, and recognizing what kind of mistake has been made, NAO can already adapt the instructions and help improve learning of the student immediately.

Improvements can be made by elaborating the feedback. Providing only correct or wrong to an answer might not help the student in improving his or her math skills. NAO could provide elaborated feedback immediately after the answer is given by the student.

2.3. Robots in Education

A lot of research has been done on robots in education. Most research is in line with modern pedagogical theories of learning, like Papert's constructionism theory [27], active learning [11], learning by design [10] and social constructivism as proposed by Vygotsky[41].

Robot-based learning of geometric primitives using the TURTLE/LOGO platform was introduced in the 80s by Papert, using his theory of constructionism. Papert concluded that student-robot interaction improves the student's knowledge and problem-solving skills because of an increase in motivation by students [15, 27].

Mubin et al. [24] have provided an overview of post-2000 research papers on robots in education, varying from programming a robot, teaching sciences, or teaching a second language. In his review, all sorts of robots pass by. Each robot has been carefully selected, to fulfill the task as good as possible. For example, an electronic robotic kit has been used to teach the basics of electronics. A LEGO mindstorms robot (mechanical robot) uses sensors to explain gravity to students, and a humanoid robot to teach English or understand physics by kicking a ball.

Janssen et al. [14] have done research using a humanoid robot teaching mathematics to students in the age of 9 and 10. Their main focus was long-term motivation. They found out that an imitation game, in combination with different levels of arithmetic tasks provided an increase in motivation.

2.3.1. Robot Behavior

The robot can take on a number of different roles in the learning process, with varying levels of involvement of the robot in the learning task. The choice depends on the content, the instructor, the type of student and the nature of the learning activity[24].

In general, younger children like to work with robots. They expect robots to take the role of a private tutor rather than a learning companion [31, 32]. NAO will take the role of a tutor, and thus should behave as one.

Kennedy et al. have done research in the field of a social math tutor robot to see if a social robot improves learning. They conclude that up until a certain level of social behavior, the student will learn significantly more, but above that level, the student will get distracted easily [16]. A robot should behave humanlike, in a way of gazing, and showing interests by slow arm gestures[13].

2.4. Gamification

Gamification is to use elements of game design in non-game contexts, products, and services to motivate desired behaviors [8]. Gamification has been seen as a successful part of education, as it motivates children to perform better [17, 18]. Snappet and De Rekentuin also make use of gamification, as part of their educational tool. Core difference in using a robot instead of a tablet, is that the robot has no element of view, while the tablets can show a lot. Can we add an element of gamification to NAO, to help improve the learning results?

According to Kiryakova, educational content for gamification should allow for the following:

- Multiple performances: Students should be able to repeat the same exercises again, in case of an unsuccessful attempt.
- Feasibility: Exercises should be achievable, they should be adapted to students' potential and skill level.
- Increasing difficulty: If exercises were to be completed, more complex exercises should be provided.
- Multiple paths: Exercises should be solvable in different ways, for students to develop diverse skills, and their own strategy to solve an exercise [18].

According to Kickmeier et al.[17], adding a competitive element to a learning objective can already increase the learning results. NAO does not make use of any tablet, or visual output, and therefore is limited to movement and speech.

2.5. Feedback

The term feedback is often used and in daily speaking can be any type of advice, praise, and evaluation. Feedback is a consequence of performance, as feedback always happens after a response. Actual feedback, however, is information about how we are doing in our efforts to reach our goals. It is part of the teaching process, but has no effect in a vacuum. For feedback to be powerful, there needs to be a learning context to which the given feedback is addressed.

Hattie and Timperley provide a model of three questions in using feedback:

1. Where am I going? - What is the goal of the feedback? A critical aspect of feedback is the information given to a student about the learning goals of the assignment.
2. How am I going? - Am I on track regarding the goal? Feedback is effective when it consists of information about progress and how to proceed.
3. Where to next? - What to do next? Giving information about the next step can have some of the most powerful impacts on learning.

The answer to these questions enhance learning when there is a discrepancy between what is understood and what is aimed to be understood. It can increase effort, motivation, or engagement to reduce this discrepancy. High motivation can be seen as a prediction to good results [38]. These feedback questions should be intervened with the focus of the feedback. Which can be done on 4 levels:

1. Feedback about a task or product. - Direct feedback about the product: what is correct or incorrect, what should be added or removed.
2. Feedback about the process of a task. - Explaining why the process is wrong or should be adapted.
3. Self-regulation feedback. - Give feedback to apply already known knowledge into the new assignment.
4. Feedback directed to the self. - Anything to praise the student, like ”Well done”, or ”You are a great student”.

According to Hattie and Timperley, feedback directed to the self is the least effective, while self regulated feedback and feedback about the process of a task are most powerful in terms of mastery of tasks [12, 25, 40, 43].

According to Wiggings [43], research supports the idea that feedback improves our way of teaching, and is better than teaching on its own. This means that providing problems for students to solve and give peer instruction to each other provides frequent and continuous feedback about the level of understanding the matter. Wiggings agrees with Hattie and Timperley that information only becomes feedback if the information tells something about where you are now, and if that’s okay or you need to change track.

Children need tangible feedback. That is how they learn to walk, hold a spoon, understand that some words yield food or drinks. Even if the feedback is specific and accurate, it is not of much value if the user is overwhelmed by it, or does not understand the given feedback. In most cases, the sooner the feedback is given, the better. Feedback should be timely instead of immediate, because immediate feedback can also overwhelm the student. Badly timed feedback is a big problem in current education. Vital feedback on key performances often comes too late, weeks after the performance. Wiggings also mentions that technology is a powerful tool in providing feedback, as it can provide unlimited, timely feedback and opportunities to use it. Like Wiggins, Singh has done research in the field of immediate feedback and concludes that computer-based homework with immediate feedback is better for students compared to traditional homework [33].

Verschuren has done quite some research among students and primary schools about feedback. He states that a big positive influence on students can be:

- Accelerating the matter. Challenge students and formulate a clear goal in the learning situation. Teachers often put the bar too low for students, as they are afraid of overcharging the students with questions.

- Give feedback on the result as well as the road to the result. At schools, students often receive just mere seconds of feedback, which is not even focussed on the process.
- State a clear road to success. Students need to know when they are doing okay, or need to change track.

Research among students from 7th and 8th grade (age 10 to 12) also states that students do not feel challenged at all, and do not think math problems are hard to solve. Creating their own goals in learning works very motivating for students, but they need to be guided in this process [40].

2.6. Conclusion

When math education by teachers starts, children have already gained some incidental knowledge doing simple calculations in everyday situations. When children enter primary school, they start structurally learning math following objectives defined by SLO. This research focuses on the objective of adding and subtracting number up and until 100, which is mostly trained in 4th grade (age 7 and 8). This research will therefore focus on the 5th grade (age 8 and 9) to see if students memorise, understand and apply the taught matter, with help of NAO. In 4th and 5th grade, students use and repeat various simple strategies and calculation strategies. The generally most frequently used strategies are lacing and splitting. Therefore, NAO will focus on these two methods. NAO also aims to identify commonly made mistakes such as switching units in bigger number or confusion about symbols or strategies. Solutions to these problems include a lot of practice, immediate feedback, explain strategies and cheering. NAO will aim to provide an environment without stress and distraction.

Current teaching aids include Snappet or De Rekentuin. The aids both use tablets or computers for students to give their input. NAO is different from these tools in adapting instructions and giving elaborate feedback independently of the teacher.

NAO, as a teaching robot, can offer improved knowledge and problem-solving skills as well as increased long-term motivation when they are used in educational settings. In general, young children seem to like working with robots, as they are expected to act as a private tutor rather than as a tool. A robot should behave socially up until a certain level to stimulate learning but prevent distraction. It should behave humanlike with arm gestures. Furthermore, like Snappet or De Rekentuin, NAO could add an element of gamification by adding elements of game design to non-game contexts. In that way, students seem to perform better. The gamification elements should allow the student to repeat exercises in case of unsuccessful attempts, as well as it should include feasibility, increasing difficulty and multiple possible paths.

An important element of this research is feedback, which concerns information about how we perform in efforts to reach goals. It requires information about the goal, the followed track and the next steps. Feedback seems to be most effective when it is self-regulated and intended for process. Children need tangible and timely feedback, not too early (to prevent overwhelming the student) and not too late. Giving feedback on the result as well as on the track, challenging students and defining a clear road to success are crucial in helping the student succeed. NAO should give feedback according to an answer of the student, and adapt the feedback according to that answer.

3 Design

The second step in answering the research questions, is designing NAO. The goal of the project is to be able to use NAO as a teaching application in math for children. NAO should give feedback according to an answer of the student, and adapt the feedback according to that answer. In this chapter, some personas and scenarios will be stated. These scenarios are completely described in detail in sections 3.2 and 3.3, providing for use cases in section 3.4. These use cases help implement NAO to provide good feedback. The interaction of NAO with the student and design choices made are explained next.

3.1. Personas

In order to create scenarios and use cases for NAO to work with, we need to define some personas. These personas are made up examples of the real life situation, and their mere use is to clear up the scenario or use case.

NAO Robot: Pixel is the NAO robot used to do math with students in primary schools.

Student: Anna is 8 years old, she just started in 5th grade of primary school. She lives near Delft, in a fairly crowded city. She grew up with both her parents in a small family. At school she prefers to play outside instead of making her assignments, but she manages to score average results.

Student: Adam is a 9 years old boy. Adam has done 5th grade last year, but was not good enough at math to proceed to 6th grade. He lives in a small village near Delft and works hard to go to university when he is older.

Teacher: is 31 years old, and loves to see innovations in education. Teacher likes the idea of getting analytics from tablets and applications as Snappet and Rekentuin and is curious what NAO can add to the educational system.

3.2. Scenario 1: Meeting the robot

Anna and Adam are in the same fifth grade. They have both worked with Snappet on a tablet before, to play and learn math. Anna heard that a robot was coming to the classroom to help with math. Anna is excited and a bit afraid as she has never seen a robot before, except in movies. Adam has played with a small robot dog before, and is not afraid to work with the robot to increase his math level. Adam is a bit anxious about what the robot would look like though. When Pixel enters the room, the students are very curious to see what it can do. Pixel starts off talking about himself, and showing some dance moves. The students are amused and look forward to working with Pixel.

3.2.1. Interaction Flow

A lesson plan usually involves an initial introduction (introduction phase) and a phase in which students apply their knowledge practically (intensive phase). The introduction phase usually helps when the student or educational instance is unfamiliar with the use of robots in education [24, 29]. Therefore, in order to successfully start the sessions, a class-wide introduction session of the robot should be made, where the robot is introduced to the students, and can show what it is able to do. During this class, Pixel will introduce himself, show some dance moves to loosen up the mood, and asks the children to come close and have a look at him from nearby. They can touch him, lift him op slightly, after which Pixel will tell them that he hopes to help them with math in the next few days.

In figure 3.1 one can see the entire flow diagram of the first meeting. This is mostly automated, but under supervision of the researcher and the teacher.

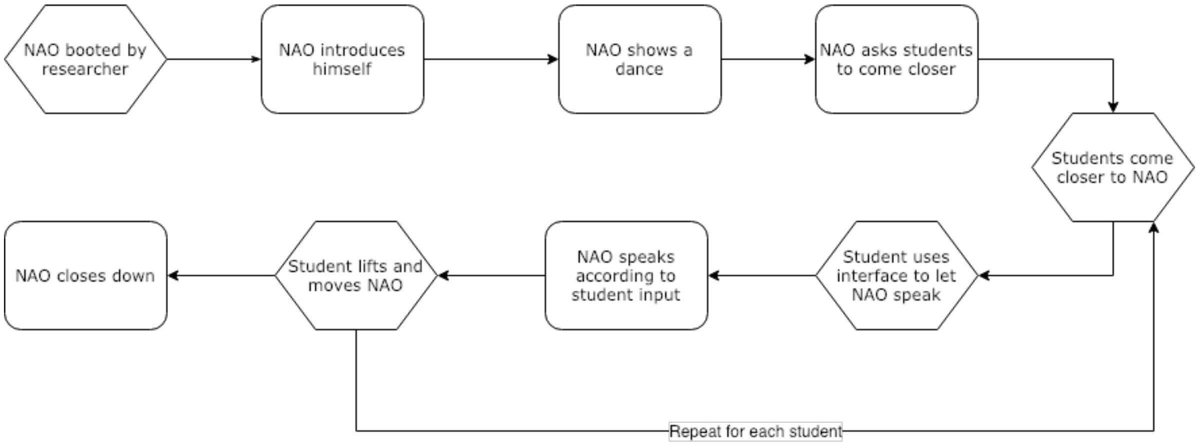


Figure 3.1: The introductory class, described in a flow chart.

Because this scenario is only an introductory class, no use cases will be created. Focus will be put on scenario 2: Math tutoring one on one.

3.3. Scenario 2: Math tutoring one-on-one - Session with NAO

Pixel is standing at the front of the table, facing towards the empty chair, when one of the students enters. Pixel welcomes the student, after which it will explain what is going to happen. Pixel will give the student an exercise to answer. When the answer is correct, the student will receive a motivating reward. When the student answers wrongly, Pixel will give feedback to the student, and lets him or her try again. After 20 minutes of practising, Pixel will thank the student for his or her help and attention, and will go back to rest.

3.3.1. Interaction Flow

A teacher which first explains a subject and then engages in a dialog with the student proved to be more effective. [31, 35] When NAO starts the one-on-one session, it will therefore first explain what is going to happen: NAO will tell the student that they learned addition and subtraction of numbers up and until 100 last year. NAO will do math with the student twice, for 20 minutes per session. Whenever a student correctly answers an exercise, NAO will reward the student. Whenever a student wrongly answers an exercise, NAO will try to provide adapted feedback to the student, in order to help the student to perform better next time. Adapted feedback in this context, means feedback according to the answer provided by the student. After the feedback, NAO lets the student answer the same exercise again. If answered wrongly the second time as well, NAO will explain the exercise and how to get to a good answer, else it will reward the student and continues with the next exercise. After the 20 minutes session, NAO closes down by either telling the student that it will see him/her again next time, or by telling that the student worked very hard, and hopes to see him or her again sometime. At the end of the session, NAO will always perform a dance to motivate the student for the next session, and let the student go back to the classroom with a positive feeling.

An entire session with NAO can be seen in the flow diagram in figure 3.2

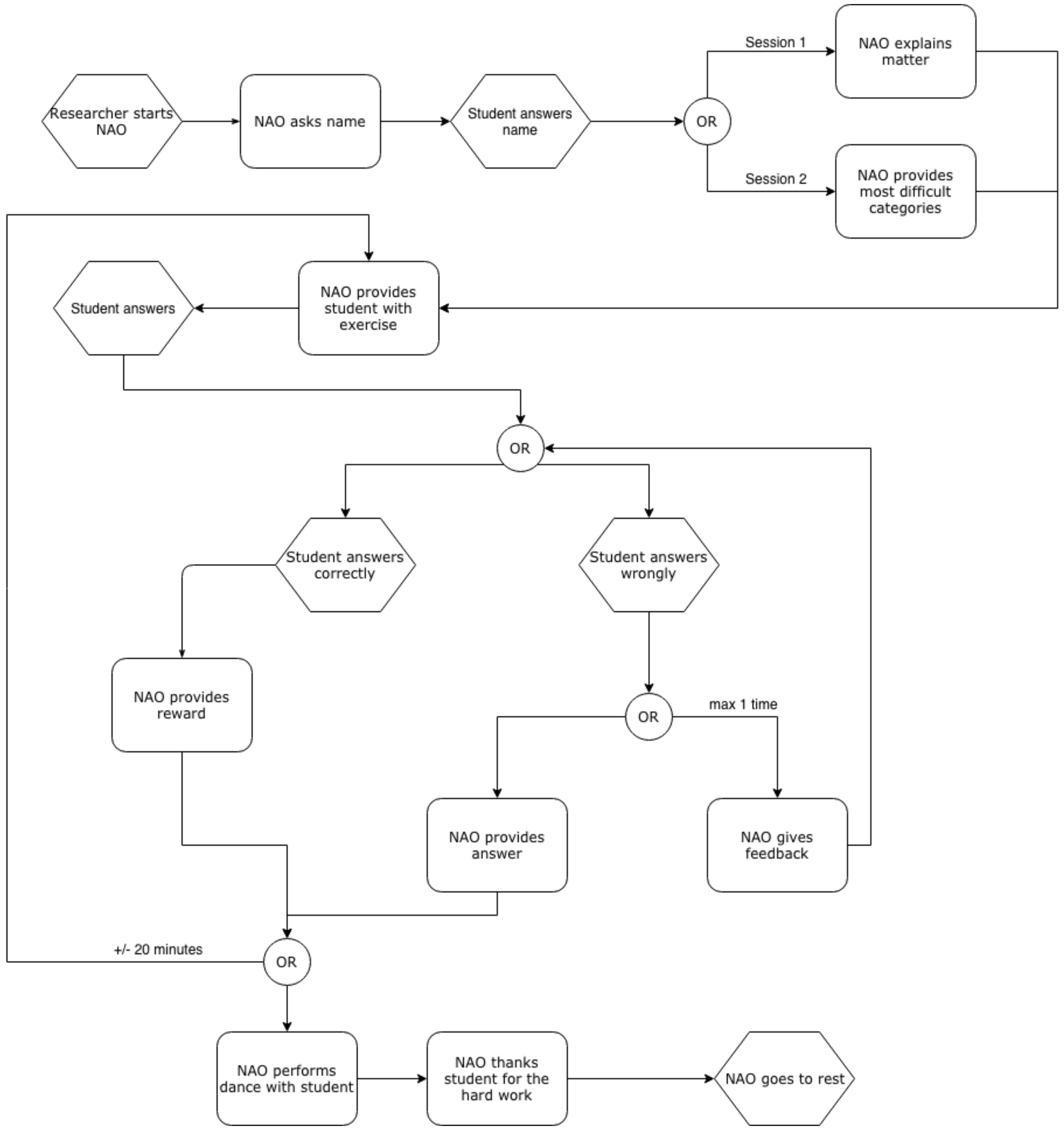


Figure 3.2: A session with NAO, described in a flow chart.

The figure in 3.2 is the flow diagram for the group that receives elaborated feedback. As stated before, two groups will be created. A test group, who will receive elaborated feedback from NAO, and a control group, who will receive simple feedback from NAO. The difference for the group that does not get elaborated feedback, is that NAO will not provide an explanation of the exercise after two wrong answers, and the feedback is minimal (e.g. ”wrong” or ”correct”).

3.3.2. Exercises

During the session, the student will be provided with exercises for about 20 minutes. The exercises provided are split into 7 categories:

1. Passing tens *(Dutch: ”Over het tiental heen”)*.
2. Adding tens and units *(Dutch: ”Tientallen plus eenheden erbij”)*.
3. Adding tens and units, passing tens *(Dutch: ”Tientallen plus eenheden erbij, over het tiental heen”)*.

4. Through tens *(Dutch: "Door het tiental heen")*.
5. Remove tens and add later *(Dutch: "Tiental wegdenken en er later weer bijplaatsen")*.
6. Tens minus tens and units *(Dutch: "Tientallen min tientallen en eenheden")*.
7. Tens and units, through tens *(Dutch: "Tientallen en eenheden, door het tiental heen")*.

NAO will provide exercises one by one, for each category. So first one from passing tens, the next exercise from tens and units, and so on. In the math tests provided to the students as described in chapter 5 a subcategory is used: turning around *(Dutch: "Omdraaien")*, where the student should add the smaller number to the bigger number. This is actually the same as tens and units, passing tens, and therefore not implemented as a category. The entire list of exercises can be found in appendix L.

3.4. Use Cases

To provide a step-by-step solution of what NAO should be able to do, several main use cases have been created. These use cases are used to describe exactly what happens in every situation during the session from section 3.3. These use cases will provide the specifications for the implementation in chapter 4. Every use case consists of the same data: a pre- and post-condition, an action sequence, requirements in order to successfully run the use case and claims that should be reached by running the use case.

3.4.1. UC01.1: NAO starts: Asking the student's name

The main purpose of this use case is for NAO to ask the student's name. This is only done in the first session. NAO will not use the name of the student in further conversation, because of possible mispronunciations of the name. This might cause laughter, giggling or frustration for the student and distract him or her from the main purpose: practising math. When the child provides a name to NAO, the student will get the feeling that the conversations get more personal. The student will then hopefully have a positive feeling, knowing that NAO is working one on one with the student. The basic requirement is the speech recognition of NAO. The claim that is tried to reach, is to improve the session by providing a personalised conversation.

ID	UC01.1
Objective	Objective is to create user data and personalise the conversation.
Actors	Student, NAO
Pre-condition	Student is sitting in front of NAO, NAO has booted.
Post-condition	Student gave NAO its name.
Action sequence	1. NAO opens conversation. 2. NAO asks for student's name. 3. Student answers with its name. 4. NAO welcomes student, explains what is going to happen.
Requirements	1. Speech recognition of NAO is turned on.
Claims	1. Personalised conversations improve the session.

3.4.2. UC01.2: NAO starts: Getting child data

In the second session, categories that went wrong in the first session will be provided to the student. In order to do that, NAO will need to load data from the previous session. This data is already saved to the filesystem, so the researcher will start the script according to the student number, after which NAO will load the top three categories in which the most mistakes were made. The claim that is tried to reach is to adapt the second user session according to the data of the first session.

ID	UC01.2
Objective	Objective is to load previous data from the child, in order to continue monitoring progress.
Actors	Student, NAO
Pre-condition	Student is sitting in front of NAO.
Post-condition	Student is ready to start with the exercises.
Action sequence	1. NAO opens conversation. 2. NAO loads exercises from categories that went wrong the first time. 3. NAO welcomes student, explains what is going to happen.
Requirements	1. Student has done the first session. 2. NAO is provided with session information.
Claims	1. Previous data can be used to adapt second user session.

3.4.3. UC02: NAO gives an exercise to the student

The main purpose of this use case is to provide an exercise to the student to solve. Before starting exercises, NAO provides the student with information about the taught matter, to remind him or her what was taught last year. The exercises belonging to all the categories are loaded by NAO, and one by one, an exercise from each category will be provided to the student. The touch sensors in the head and feet are used to either give an answer or repeat the question, and the speech recognition is used to hear the answer. The claim that is tried to reach, is increasing mathematical skills by providing exercises in the taught matter.

ID	UC02
Objective	Objective is to provide the student with an excercise so that (s)he can practise math.
Actors	Student, NAO
Pre-condition	Student is sitting in front of NAO. NAO loaded exercises.
Post-condition	Student has given an answer to NAO.
Action sequence	1. NAO explains the matter which it is going to practise with the student. 2. NAO provides an exercise to the student. (a) Student did not hear the exercise. Student touches NAO's head. Repeat 2. (b) Student did hear the exercise. Student touches NAO's foot. Continue with 3 3. Student answers the exercise.
Requirements	1. Use speech recognition. 2. Use touch sensors.
Claims	1. Student can practise the taught matter in order to increase his/her mathematical skills.

3.4.4. UC03: Student gives wrong answer: NAO gives feedback

The main purpose of this use case is providing the student with relevant feedback. In order to give relevant feedback, NAO must detect a mistake that has been made by the student. The feedback that will be generated according to the answer given should help the student in learning what went wrong in the calculation process. In case the wrong answer could not be detected as one of the mistakes, NAO will provide simple feedback. More about feedback and detecting mistakes can be found in sections 3.6 and 4.3. Speech recognition is used to hear the answer. Claims that are tried to reach, are NAO who is able to detect mistakes according to a given answer, and students should understand why mistakes are made.

ID	UC03
Objective	Objective is to give feedback to a student, so that it can learn from its mistake.
Actors	Student, NAO
Pre-condition	Student is sitting in front of NAO and has given an answer to NAO's exercise.
Post-condition	Student is provided with feedback.
Action sequence	1. NAO checks answer: answer is wrong. 2. NAO tells child that the answer is wrong. (a) Mistake detected: NAO provides adapted feedback according to the answer given. (b) No mistake detected: NAO simply tells student the answer was wrong. 3. NAO tells student to try again.
Requirements	1. Use speech recognition. 2. Detect mistake. 3. Generate feedback.
Claims	1. NAO can detect mistakes according to answer given. 2. Student understands why (s)he made a mistake.

3.4.5. UC04: Student gives correct answer: NAO gives reward

The main purpose of this use case is providing rewards to the student. By rewarding the students, each student should stay motivated to work hard and give correct answers. Rewarding is done by using speech, movements and LEDs from NAO. More about rewarding can be found in sections 3.7 and 4.4. The claim that is tried to be reached, is that students will remain motivated because of the reward.

ID	UC04
Objective	Objective is to reward the student to keep motivation high.
Actors	Student, NAO
Pre-condition	Student is sitting in front of NAO and has given an answer to NAO's exercise.
Post-condition	Student has gotten a reward from NAO.
Action sequence	1. NAO checks answer: answer is correct. 2. NAO praises the student. 3. NAO gives student the reward.
Requirements	1. Use speech recognition. 2. Use LEDs in eyes.
Claims	1. Student has higher motivation because of reward.

3.4.6. UC05: NAO provides explanation to the student

The main purpose of this use case is to provide an explanation of the exercise to the student. The student has received feedback on the first given answer to the exercise, after which he or she should be able to get to a correct answer the second time. When the student has given a second wrong answer, NAO must explain how to get to a correct answer, so that the student knows how to solve calculations in the category provided. A student only has two tries per exercise. This is done, because a student can get lost in an exercise for too long. As similar exercises in the same category will be provided to the student, it can practise the matter again. To provide an explanation, NAO needs to know explanations of exercises from each category. The claim that is tried to be reached is that the student understands the explanation.

ID	UC05
Objective	Objective is to provide explanation and answer to the student, for him/her to see how to get to a correct answer.
Actors	Student, NAO
Pre-condition	Student is sitting in front of NAO and has given two wrong answers to NAO's exercise.
Post-condition	Student has gotten explanation and answer from NAO.
Action sequence	1. NAO tells student the answer is again wrong. 2. NAO will explain how it was taught him, by providing a step-by-step solution.
Requirements	1. NAO knows the explanation of an exercise.
Claims	1. Student understands the explanation.

3.4.7. UC06: NAO saves progress

The main purpose of this use case is saving the progress of the student. Every exercise and the amount of tries needed is saved to a file, to provide information about the used categories in the second session. In order to save the progress, NAO needs access to the progress of a student. The claim that is tried to be reached is that saved data can be used for future purposes by either NAO or teachers.

ID	UC06
Objective	Objective is to save data according to provide teacher with information and provide adapted feedback in the second session.
Actors	Student, NAO
Pre-condition	Student has finished an exercise.
Post-condition	NAO saved progress.
Action sequence	1. NAO saves progress to list. List will be saved to file after the session.
Requirements	1. NAO has access to the progress of a student.
Claims	1. The data saved can be used for future purposes by teachers and NAO.

3.4.8. UC07: Student is finished, closing down

The main purpose of this use case is to close down NAO after a session has ended. The student will be praised for his/her hard work and will receive a dance as reward for working so well. Afterwards, NAO saves the progress to the file system. In order to do this, NAO needs to know how to perform the dance, and needs to have writing access to the file system. The claim that is tried to be reached is that the student enjoyed working with NAO, and wants to do this again.

ID	UC07
Objective	Objective is to close down in a good way, for student to want to do this again.
Actors	Student, NAO
Pre-condition	Student finished all exercises.
Post-condition	NAO is in resting position, student left.
Action sequence	1. NAO praises student for the hard work. 2. NAO lets student know how many exercises it had done correct. 3. NAO performs the dance. 4. NAO saves list with student data to file. 5. NAO says goodbye and goes in resting position.
Requirements	1. Perform dance moves. 2. Write to filesystem.
Claims	1. Student enjoyed working with NAO, and wants to do this again.

3.5. Interaction Design

Applications like Snappet and De Rekentuin have been proven effective in teaching math to children. However, research done by Van Gorp as mentioned in the Belgian newspaper De Standaard, proves that more calculation mistakes are made when children work with tablets instead of pen and paper. Beside that, it is proven that tablets are mostly seen as toys, therefore children get distracted easily when working on tablets or they do not take the exercises seriously, so one more mistake does not matter [1].

Next to distraction, cognitive load theory as described by Sweller [34], can also be part of the problem. Using a robot, a tablet, speech and a new environment for the child to work in, could overload the brain of a child and thus have a negative effect on learning math.

Saerbeck et al. [31] noted that a tutor is closer to the student and can explain the learning content better for a student's needs. Bonding is important as well, and can easily be done by using ”we” instead of ”you”, and using motivational sentences such as ”Don't worry, I will help you”, or ”It was not easy for me either”. NAO is able to perform this role better than a tablet, therefore, no tablet will be used.

Batliner et. al. [7] have done research about how children react to a robot who does not listen, or completely ignores what the children say. They state that children are likely to change strategies if the machine does not understand them. Children repeat, reformulate, use other words or simply use a pronounced, marked speaking style. If the robot keeps on ignoring or not understanding what the child says, he or she most likely will get angry or impatient. Therefore a training session will be developed, which teaches the students how to talk to and interact with NAO, in order to let NAO understand the numbers which the children say. This training session can be found in appendix A.

3.6. Feedback Design

Tiberius and Billson [35] analysed social supportive behavior in context of education. They compare two ways of interaction between a robot and student: 1. teaching as knowledge transfer and 2. teaching as a social dialogue. They argue that a dialog is more appropriate, hence more effective. NAO will therefore use a dialog in order to provide the student with feedback.

Feedback is also influenced by the way NAO uses non-verbal behavior. Non-verbal behavior can be used to guide the student's attention. For example to control the flow of the dialog, or give feedback. One big advantage of non-verbal behavior is that using this way of feedback does not disrupt the student's way of thinking. E.g. a nod or shaking the head can be used to steer the student in the right direction [31].

As mentioned in section 2.5, a model of feedback according to three questions is an effective way of providing feedback. As NAO is providing feedback on the process of a task, the provided feedback will answer the three questions:

- Where am I going: The goal of the feedback is to help the child realise what it did wrong.
- How am I going: NAO will provide feedback on the step that most likely went wrong, mentioning that a correct result might be found if the exercise is redone.
- Where to next? Try the exercise again, if the directed feedback did not work, NAO will explain how to get to a good answer for the next exercises to be made.

The version of NAO with elaborated feedback is programmed according to the model of Hattie and Timperley. The version of NAO with simplified feedback, only mentions whether the answer was wrong or right.

In case the answer is correct, positive feedback will be given. This will be the same for the test group who receives elaborated feedback, as for the control group, who receives simple feedback. This positive feedback comes along with a reward, as described in the next section.

3.7. Rewards Design

As stated before, motivation and learning go hand in hand. Therefore, applying gamification or 'learn by playing', we need to create a reward for the students, which motivates them to continue, as well as makes them eager to learn and memorise the calculation steps. Janssen et al.[14] have used an imitation game to motivate children to continue working. A good and fun way to memorise steps and keep high motivation is dancing [9].

Rewards will be given to the students whenever they succeed in completing an exercise provided by NAO, and the answer is correct.

Rewards in between the exercises will be given by cheering. The LEDs in the eyes of NAO will turn into a rotating rainbow colour, and it will cheer some motivating and complementing words. Variations are made by changing the words that NAO says to the student. The variations of cheering and motivating after a correct answer are:

- Great job! That indeed is the correct answer *(Dutch: "Goed gedaan! Dat is inderdaad het goede antwoord!")*
- Correct again! You're doing great! *(Dutch: "Alweer goed! Je bent goed bezig!")*
- You're the best! That is the correct answer! *(Dutch: "Je bent een kanjer! Dat is het goede antwoord!")*
- Yes, correct! Let's go for the next one! (*Dutch: "Ja, helemaal goed! Op naar de volgende!")*

A big last reward will be a short dance. NAO will teach the students how to dance the Tsjoe Tsjoe Wa (MiniDisco - Tsjoe Tsjoe Wa. A video performance can be found here: `https://www.youtube.com/watch?v=XveIiC22DYI`). This easy to learn dance consists of 9 simple steps that repeat themselves over and over again. It is a dance that most children are familiar with as it is often performed at places like campsites. After completing the 20 minutes session with NAO, it will provide the dance to the students, and asks to dance along.

As the rewards are used as a motivational impact, and are not involved in the feedback mechanism, both groups will be provided with rewards.

3.8. Conclusion

To conclude, this chapter about the design of NAO started by stating personas to work with during our scenarios. Two scenarios were formulated; an introductory class to let the children interact and touch NAO, and a math tutoring session, as will be used during the experiment. Use cases were created, stating the most important parts of NAO to be implemented in order to create a successful math session. The design was ended by specifying interaction, feedback, and rewards. Interaction is specified by only using speech and movements. Feedback is specified as answers to the model of Hattie and Timperley. Rewards are specified by using motivational texts, rainbow LEDs in NAO's eyes and a dance reward at the end of the session, to keep students motivated to increase their learning effect.

4

Implementation

In the previous chapter, the way NAO will be designed is stated. This chapter explains in detail the implementation of this design: the implementation of NAO. Which choices have been made during implementation, which problems have arisen, how were they solved or worked around? Continuously looking back at how the design was initially planned to be, differences will be explained. First, the structure of the code will be explained, after which the interaction and behavior will be described. In the feedback section, the implementation of the feedback by NAO is described. Next, the rewards section describes the implementation of the motivational rewards that have been given by NAO. Lastly, all implemented parts need to be tested, to see if NAO is ready for the experiment.

4.1. Structure

The structure of the code is split up into three Python packages: a training package, a feedback package and a non-feedback package. Each package is created and used for a certain part of the experiment. This chapter will explain what function each package has and how this is implemented.

4.1.1. Architecture

The three packages as named earlier all depend on NAOqi, which is a built-in package by Aldebran. In figure 4.1, the architecture of the packages in NAO is displayed.

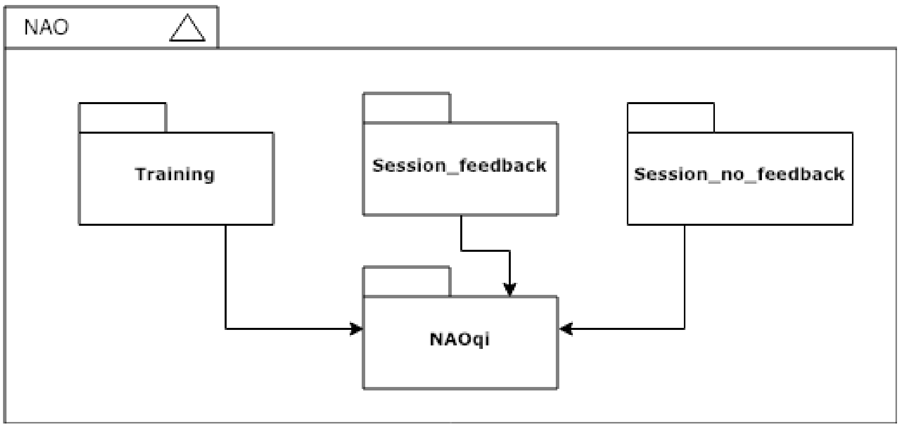


Figure 4.1: Architecture of packages installed on NAO.

4.1.2. Training Package

The training package is used for a special training for the student. During the training, the student will learn how NAO works, how the student needs to talk to NAO in order to be understood, and where the buttons are located in order to function without problems during the session.

The training session consists of three parts:

1. Repeating numbers,
2. Touching NAO,
3. Solve easy, basic math problems.

Repeating a number is implemented by using the speech recognition of NAO. NAO calls a number, after which his eyes turn blue and speech recognition is activated. The student has three seconds to repeat the number, in order for NAO to understand him. Afterwards, NAO will tell the student what it understood and when this is the same number, it continues. If not, the student must try again, until the correct number is understood. This way, the student learns how to pronounce certain numbers in a way that NAO can understand, to make sure less mistakes are made during the session.

Touching NAO is implemented using a ”Tickle Game”. A Tickle Game is a game in which a student has to tickle NAO. Touching NAO at certain body parts (read: front of the feet and top of the head), leads to a laughing sound, produced by NAO. Each body part has a different laughter sound, to keep some variation in the game. The Tickle Game is done to learn the students where the touch buttons of NAO are located in order to activate speech recognition or repeat a question during the math session.

Last but not least, a session-like part is built into the training package. This is implemented the same way the session is, but only using three basic, simple exercises to be solved. This is done to minimise the mistakes that the student can make in the exercise, and in the meantime is able to test the entire workflow of an exercise during the math session.

An entire listing of the package files and it’s methods can be found in appendix C. The full script, in which the exact conversation is typed out, can be found in Appendix A.

4.1.3. Feedback Package

The feedback package contains the session with feedback, given to the student after (s)he has given an answer to an exercise. This can either be a correct or wrong answer. The package is built up from the steps as described in sections 3.3 and 3.4

From the beginning onward, NAOqi is one of the most important packages that have been used while implementing the session. Speech recognition, autonomous life, movements, text-to-speech, are some of the functions that NAO has built in and widely used during the session. The session starts off asking for a name, to make a personal impression, after which it will explain what has been taught before and what it is going to do for the next 20 minutes. NAO will not use the name of the child during further interaction, however. This is done because of the way the text-to-speech is programmed. As we live in a multicultural society, with names that are not seen very often, it might be hard for NAO to pronounce the name correctly. A mispronounced name, might lead to frustration or laughter, which will cause a loss in concentration. Therefore names are kept out of the rest of the conversation. Afterwards, NAO will start the math session of 20 minutes using a timer. All the exercises are saved in a json file, to be loaded when NAO starts. This json file can be found in appendix L. One by one, these exercises are given to the student to solve. When the student knows the answer, he or she presses the front of the foot of NAO to activate speech recognition. When the student would like to hear the exercise again, he or she can touch the top of NAO’s head. More about the challenges of interaction between NAO and the student can be found in section 4.2.

When the student has answered correctly, NAO will praise the student as described in section 3.7. More about rewards can be found in section 4.4. When the student has answered wrongly, NAO will provide feedback as described in section 3.6 and appendix B. More about the implementation of the feedback can be found in section 4.3. In case the answer is wrong twice, NAO will explain the exercise as a tutor, explaining as if they are calculating together, as described in section 3.5.

Every exercise done is saved to a list. This will make sure that NAO is able to tell the student in the end how many exercises it completed correctly at once. Also, it is needed for the second session in order to know which categories to focus on. The focus of the categories is calculated by picking the three categories in which the most mistakes are made. Especially, NAO is looking at categories in which the student gave a wrong answer to an exercise twice.

When the 20 minutes are over, NAO tells the student how many exercises were correct at once, thanks the student for the hard work and performs the dance as described in section 3.7.

The entire list of files and methods used in the feedback package can be found in appendix C. The full script, in which the exact conversation is typed out, can be found in appendix B.

4.1.4. Non-feedback Package

The non-feedback package consists of the same algorithms and functionality as the feedback-package described above. The only difference here, is that instead of helping the student to improve, NAO will simply state whether the answer is wrong or correct.

The entire list of files and methods used in the non-feedback package can be found in appendix C. The full script, in which the exact conversation is typed out, can be found in appendix B.

4.2. Behavioral Challenges

As written in section 3.5, NAO will play the role of a tutor and cheer informally as if it had to do the learning phase himself as well.

4.2.1. Autonomous Life

In order to make NAO look alive, the ALAnimatedSpeech has been used when NAO is talking to the student. This proxy lets NAO make normal hand gestures when it is speaking. NAO has a built-in Autonomous Life functionality. This functionality is on by default, and provides a ”deaf and dumb”, feeling alive movement sequence. Deaf and dumb here, means nothing more than making movements as if it is waiting for something to happen. This is applied whenever NAO is waiting for the student to answer a math problem.

Another method to make NAO look alive, is tracking. We humans, look at someone when we talk to them. NAO is able to track faces using the ALTracking proxy or the basic awareness proxy. First, basic awareness was implemented. A person was easily tracked, and followed when moving left, right, up and down. However, when the person reached out of sight, a problem arose. When this happened, NAO would make weird movements with his head, continuously looking at the ceiling, a random head movement from entirely left to entirely right at maximum speed, or other random head movements. This felt very unnatural, and when this was replaced by the ALTracker proxy, the problem was solved, as such that losing a person didn’t lead to weird movements anymore. When NAO loses sight of a person now, its head will stay still, until it sees something that looks like a face. In that case it will move the head accordingly and start following the person again. This feels more natural, and therefore used in the final codebase.

4.2.2. Interaction with NAO

A challenging problem is NAO’s speech recognition. Words and numbers are misunderstood quite often. Where the confidence level started at 0.6, after some test runs with fellow students, a confidence level of 0.5 seemed to work quite well. After the pilot, which will be described in section 4.5.2, this confidence level was adjusted slightly to 0.45, as most students reached 0.47. More about the procedure of the pilot can be found in section 4.5.2 The training as implemented (section 4.1.2) will provide a practise run for students, to repeat five numbers, after which NAO will tell them what it heard. This way, the students learn in what way they have to communicate with NAO. It has been noticed that the speech recognition is focused on the point of view of NAO. The speech recognition works better, when NAO is looking at the talking person. Relating to the head problem of basic awareness, whenever NAO might lose the student out of sight, the head stays still and directed in the direction of the child. This improves the speech recognition, even when NAO might lose track of the student in front of him.

At first the goal was to make interaction between NAO and the student as intuitive as possible. So whenever NAO asked a question, speech recognition would turn on and it would wait for the answer of the student. A problem arised here when a student is thinking out loud: the first number that NAO recognises, is being seen as the answer of the student. As students of that age are unpredictable in giving answers, the implementation was switched to a button. Before giving the answer, the student has to touch either one of NAO’s feet. Speech recognition will then turn on, after which the student has 3 seconds to give the answer. After the pilot, it was concluded that three seconds were enough for the student to give an answer, as long as it remembered the answer it calculated.

Human brains work with different kinds of memories. As the short term memory is only able to remember things for a short amount of time, and NAO is working without any visual outputs, it is useful that NAO is able to repeat an exercise during the solving process. To provide a student with repeating an exercise, it can softly touch the head of NAO, after which the exercise the student is currently solving, will be repeated.

4.3. Feedback

NAO is able to provide elaborated feedback according to the answer given by the student. At this moment, NAO is able to detect four commonly made mistakes:

- Missing the last step during calculations (e.g. 7+5 = 7+3 = 10, forgetting to add 2).
- Visualising numbers wrongly (e.g. switching 38 with 83).
- The most commonly made mistake within the split and add method. (e.g. 32-14 = 22, by first subtracting 10 from 30, then subtracting 2 from 4, because the other way around seems impossible). The complete explanation of this mistake is described in section 2.1.3.
- Using the wrong operator (e.g. 7+5=2).

Missing the last step is implemented as an algorithm which calculates the answer before the last step. When for example 7+5 needs to be calculated, the algorithm first checks the operator of the calculation (in this case "+"). It then calculates what last ten is reached, (in this case 10), by adding the numbers and subtracting the surplus of the whole number. Some other examples of missing last steps are:

- 12 + 9 = 20,
- 38 + 7 = 40, and
- 54 + 38 = 90.

Feedback on visualising the numbers is implemented as an algorithm which calculates all possible answers for switching numbers. This results in a list with a maximal length of three. In that case the following numbers are added to the list: the first number switched, the second number switched, and both numbers switched.

Some examples of lists are:

- 7 + 5 - [], Empty list, single digits cannot be switched.
- 12 + 5 - [26], 12 is switched, 5 is not.
- 7 + 35 - [60], 7 is not switched, 35 is switched.
- 63 + 18 - [54, 144, 117], first answer only 63 is switched, second answer only 18 is switched, third answer both numbers are switched.

During testing, it was found out that in some cases, the actual answer is in the list of switched numbers (e.g. 15 + 33). This is always the case when a number with double digits has the same numbers (e.g. 11, 22, 33, 44, etc.). To make sure a good answer is never seen as a faulty answer, a check is built which removes the good answer from the list of mistakes.

The split and add mistake is implemented by creating a faulty algorithm of the split and add method. If the operator is subtract, then the faulty answer is calculated by switching the numbers in the last step as described in section 2.1.3

Feedback on using the wrong operator is handled by simply calculating the answer of the exercise, using the wrong operator (e.g. 7+5 = 7-5 = 2).

For every answer the student gives, NAO will check whether the answer is one of the calculated mistakes. If one of the numbers corresponds to the answer the student provided to NAO, NAO will give feedback according to the category of the mistake. In case of missing a step, NAO will tell the student that it might have missed a step in his or her calculation. In case of wrongly visualising a number, NAO will explain to the student that it might have thought wrong about the numbers, and explains how the numbers are written down. In case of the split and add mistake, NAO will tell the student that it has made a mistake when calculating the units. In case of the wrong operator, NAO will explain which operator is used in which exercises. In case a wrong answer is given that is not in one of the mistake categories, NAO will tell the student that the answer is wrong. In all cases, NAO will provide the student the opportunity to try the exercise again.

After a second wrong answer, NAO will explain how to get to a correct answer using the lacing method. Every single step is explained, using the numbers of the exercise. The entire script where NAO mentions the feedback can be found in appendix B.

4.4. Rewards

Rewards are implemented using the LEDs from NAO's body, and variations in speech. The LEDs from NAO's body can be used using one of the proxies as described in appendix C. Using variations in speech, students are being motivated in similar, yet slightly different ways, as described in section 3.7. The motivating sentences are chosen randomly.

At the end of each session, NAO will perform a dance together with the student. The dance is implemented step by step using Choregraphe. Choregraphe is an application that provides you with a WYSIWYG, click-and-slide way to create animations, dialogs, and behaviors (`https://community.ald.softbankrobotics.com/en/resources/faq/developer/what-choregraphe`). Every step is implemented as a separate animation, to be able to perform parts of the dance. When performing the dance, all steps are being connected one after the other. Some steps in the chosen dance are physically impossible for NAO to do, like "raise your shoulders" (Schouders op) and "show your tongue" (Tong naar buiten). For the steps that are physically impossible for NAO to do, the song just tells the student what to do.

4.5. Testing

In order to test the written codebase, several tests are conducted. To test the functionality, an adult test is done. Afterwards, a pilot session is done at a daycare, to set some variables and test the functionality with students.

4.5.1. Adult Tests

In order to test the final codebase, first some adults were asked to test NAO. This is done at the TU Delft university, in a lab without other people, in order to make sure NAO's speech recognition can work correctly. The main goal of the adult tests is to see whether the instructions given by NAO are clear and the confidence level of the voice recognition is workable. Afterwards, a primary school teacher was asked to test NAO, in order to see if the explanation and choice of words was correct for children. This resulted in some changes in text, to make it easier to understand, but the instructions were clear: the adults had no problem interacting with NAO. Also, the confidence level of NAO's speech recognition was adapted to 0.5 instead of 0.6.

4.5.2. Pilot

A pilot is held at a children's daycare in Delft. Several fifth and sixth graders were asked to test the feedback session as described earlier, to validate whether the mistakes made by the students are the mistakes that are found in literature. The setup is the same way as the experiment session.

NAO is positioned in a separate room, without other people or distractions. NAO is in rest mode on a table, with a chair in front of it. The student takes place at the chair. The researcher explains that he is going to start the robot in a bit, and explains that the student will receive math problems for about 20 minutes and a dance at the end. The researcher tells the student that (s)he can leave anytime (s)he wants, and if there are any questions or uncertainties, the researcher is sitting at another table in the room to ask about it. Afterwards, the researcher thanks the student for participating and leads him or her back to the playground.

The pilot mainly resulted in a small adaption in the confidence level of NAO's speech recognition. The confidence level was adapted from 0.5 to 0.45, as most students reached a confidence of 0.47. The explanation and other textual context was clear to the students.

4.6. Conclusion

In this chapter, the structure of the code was described. It described what packages were created and with what goal, and the way the code training and sessions are implemented. Difficulties in interaction and autonomous life were mentioned as behavioral challenges. Next, the implementation of the feedback and rewards was stated. NAO will provide feedback on missing the last step, visualising number wrongly, the split and add error, and using the wrong operator. Rewards are implementated as the dance reward, and LEDs will show a rainbow. Lastly, the implementation of NAO was tested, after which the confidence level was adapted, and small textual changes were made.

5
Method

In the previous chapter, the implementation of NAO was described. In order to answer the research questions, test the hypotheses, and evaluate the effects of the feedback mechanism, an experiment is conducted. First, the design of the experiment will be explained, stating how the groups are made and what we try to test in the experiment. Then, the participants are described, and which materials are used during the experiment. Last, the procedure of the experiment is explained.

5.1. Experimental Design

Students are split into two groups: a control group, receiving minimal feedback from the robot, and a test group, receiving elaborated feedback from the robot as described in sections 3.6 and 4.3.

These groups are created according to results of a test and their overall math grade. Before the experiment starts, the students receive the selection test or pretest as can be found in appendix D. This test consists of 20 math exercises to be solved, from the categories as mentioned in section 3.3.2. This pretest is made by all students at the same time, in class. The results of the pretest in combination with the overall math grades are used to split the fifth grade in two somewhat equal groups, containing about the same amount of students with comparable results. This is done by first creating a list of all students with their overall math grade and the score of the pretest. Then, students who made an equal amount of or no mistakes are split, according to their school's math level. The school's math level in this case was given by a number from 1 to 5, or a letter A to E.

For example: Students A, B, C and D had 20 correct answers in the pretest. The overall math level of Student A is 5, of student B is 4, of Student C is 4.7, and of Student D is 3.8. In this case, student A and D are put into one group and student B and C are put into one group, to level the overall math level.

The target variable concerns the learning curve of the students, and the experiment will test the hypothesis whether or not students who received elaborated feedback will do better in the math tests. This learning curve will be tested by letting students practise math with NAO for 20 minutes. During this math session, the test group will get elaborated feedback, and the control group will get minimal feedback. After the experiment, the students made tests, similar to the pretest, to see if the feedback had any effect on their learning curve.

5.2. Participants

The experiment is held in two primary schools: OBS de Singel and De Bavinckschool. Both schools are located in Vlaardingen (The Netherlands) and each has one fifth grade, containing a total of 42 participating students. All students range in the age of 7 to 9, where most of them are 8 years old. The Human Research Ethics Committee (HREC) of TU Delft provided permission to conduct the experiment, as well as all parents/care-takers of the students.

The composition of the different student groups is relevant. Student data from both the control and test group are summarised in table 5.1 The groups are identified using the initial test as main selection criteria, ensuring the math performance of the students is about levelled as described earlier. One of the female students in the test group had to stop the experiment, because of medical reasons, remaining 41 students to divide over the two groups. As can be seen in table 5.1, the test group contains more male students then female. This is because the main variable to split was the initial test, then the grade and last the gender. For the experiment, it was more important to have two groups of similar math level than to have two groups of equally divided gender.

Group	School		Frequency
Control Group	Bavinckschool	M	5
		F	5
		Subtotal	10
	OBS De Singel	M	5
		F	6
		Subtoal	11
		Total	21
Test Group	Bavinckschool	M	6
		F	4
		Subtotal	10
	OBS De Singel	M	7
		F	3
		Subtotal	10
		Total	20

Table 5.1: Division of the test and control group.

The first group is the control group, containing 21 students (10M and 11F). The second group is the test group, containing 20 students (13M and 7F). Note that these are combined groups from the Bavinckschool and OBS de Singel. The groups are combined to make sure we are not testing one school against the other, but have students from different schools, with different teachers, backgrounds and learning environments in both groups. The test group will get feedback from NAO when an answer is wrong, and will be given a dance move as reward, when an answer is correct. The control group, will **not** receive feedback when an answer is wrong, but will receive a dance move reward when an answer is correct.

5.3. Materials

During the experiment, NAO, a laptop, and some paperwork were needed. This section will describe the exact details of all materials used.

5.3.1. NAO

NAO is a robot created by SoftBank Robotics in Paris [30]. It runs a simple Linux OS, to which one can communicate using the NAOqi API [2]. NAOqi can be used in three programming languages: C++, Python and Java, where SoftBank Robotic heavily prefers Python, as this is the programming language that has the most support. Also, most of the examples and documentation use Python as a base.

As can be seen in figure 5.1, NAO is a humanlike robot. Its length is 58 cm, and its weight is 5.6 kg. NAO has built-in face detection, using two 5 megapixels cameras. It also has 2 sonars to detect obstacles when walking, and can move all of its joints with 20 degrees of freedom. Sound is provided by 2 speakers, and NAO also has 4 microphones to use with the built-in speech recognition [3]. For this research, version 5 of NAO is used, together with OS version 2.1.4.

5.3.2. Student and Researcher Input

Students were involved in this experiment by doing tests, and filling in surveys and PANAS forms. For students, the following materials were needed:

- **Consent form**: Explains the experiment, and asked permission to enter the experiment and use the data, and permission to make pictures. To be filled in by both parents/caretakers and the student.

Figure 5.1: NAO robot. Picture from Génération Robots: https://static.generation-robots.com/5474-large_default/programmable-humanoid-nao-evolution-robot-red.jpg

- **Math test**: The test is a written test, consisting of 20 math problems, to be solved without using tools like pen and paper or a calculator. In total there were three tests to be done, a pretest, posttest and second posttest. The pretest can be found in appendix D, the first posttest in appendix E, and the second posttest in appendix F. The test results are classified as a dependent variable, in which the score is the amount of correct answers on the test. The test results have a range from [0, 20], where 0 is the minimum score and 20 is the maximum score.
- **Survey**: The survey consists of eleven questions that the student needed to answer. The survey contained questions about math in common, robots in common, and practising math with NAO. The test group received two additional questions about the feedback. The entire survey can be found in appendix G. Five of the questions were open questions, the others are on a Likert-scale of 5, where 5 is totally agree, and 1 is totally disagree. The survey was held to get a view on the students' experiences with learning and doing math with NAO.
- **PANAS**: The PANAS form is a form on which 20 emotions are listed (10 positive affections, 10 negative affections), which measures affection of a student at that specific moment [4, 42]. The form is based on a Likert-scale of 5 (5 = a lot, 1 is not at all or a tiny bit). The entire PANAS form can be found in appendix H. The PANAS results are classified as a dependent variable, in which the score is the sum of the weights. The weight is positive for a positive affection and negative for a negative affection. The PANAS results have a range from [-40, 40], where -40 is the minimum score and 40 is the maximum score.

The researcher needs the following materials:

- **Laptop**: Any laptop which is capable of opening an SSH tunnel to NAO will do.
- **Researcher form**: A form where notes during the experiment can be filled in. The researcher form can be found in appendix I.

5.4. Procedure

Before the students were able to join the experiment, students as well as their parents or caretakers needed to have a signed consent form. The consent form described the experiment, and asked permission for using the student's data in the analysis. The consent form also asked permission to take pictures.

About a week before the start of the experiment sessions, an introductory class-wide session was held. During this session, the students were able to let NAO talk, touch him, and see what NAO is capable of. The introductory class is described in section 3.2. Afterwards, students were asked to make the pretest as can be found in appendix D, in order to be able to create the groups. This introductory session, together with the pretest took about an hour.

After the groups were created and the sessions were about to start, the experiment room was set up. NAO was set on top of a table, with a chair in front of it. NAO was about arm-length away from the edge of the table, in order for the student to be able to reach it. The researcher was sitting in the back of the room, pretending to be at work, but actually keeping an eye on the experiment and the student's behavior. The student was able to see the researcher. An example of the setup can be seen in figure 5.2.

Figure 5.2: The experiment room from OBS de Singel, from the front and back.

Together with the teacher, an appointment was made that the students would tag each other when the math session is done and a new one is about to begin. At the start of the day and after lunch break, the researcher calls for the first student, after which the students will switch on their own.

When a student entered the room, the researcher told him or her to take place in front of NAO. At that moment, NAO was still in resting position.

The researcher told the student what was going to happen according to the protocol in appendix J, after which the researcher started the training session. During the training session, the researcher was sitting next to the child in order to help, guide and support him or her when needed. After the training session, the researcher needed to decide whether the student's knowledge and capabilities about using NAO were good enough to work with during the math session. If not, the training had to be done again, else the researcher asked the student to fill in the first PANAS form. This was decided according to the training score. The training score (1, 2 or 3) relates to the level of confidence the researcher has in a successful math session. The way the score was given can be seen in appendix I.

Once the PANAS form was filled in, the researcher explained the math session and reminded the student to talk loud and clear. During the session, notes were taken about the behavior and input from the student. The protocol can be found at the researcher form in appendix I. The protocol mentions information about what training scores to give, when to take notes and for what actions of the student.

After the math session, the researcher asked student to fill in the PANAS form again. Once filled in, the researcher started asking questions from the survey. The Likert scale questions were printed on a A4 paper, for the students to point the number they feel is correct. Answers to open questions were typed by the researcher at a laptop.

When all questions were answered, the researcher closed down by thanking the student and asked to send the next student. The entire first session, including the training session took about 45 minutes per student.

After all students had done the first math session, the second round could start the week after. Students were called in the same way as the first session, and the setup of the room was the same as well. When a student entered the room, the researcher explained what was going to happen according to the protocol, after which the second math session was started. Once the session was over, the student was asked to fill in the PANAS form again, followed by the survey. The researcher thanked the student for participating in the experiment, and asked to bring the next student in. The entire second math session took about 30 minutes per student.

When all students finished the second math session, the first posttest was done. The posttest can be found in appendix E. Two weeks after ending the math sessions, the second posttest was done, as can be found in appendix F.

5.5. Conclusion

To conclude, the experiment consisted of two groups: a test group and a control group, which had 20 and 21 students respectively. Besides NAO and a laptop, students were able to provide input via a math test and a PANAS test, and questions were asked using a survey at the end of each session. The researcher kept an eye on the experiment, and noted important information about the behavior of the student. Before the experiment started, an introductory class was given, after which two math sessions with NAO were done for every student.

6
Results

In the previous chapter, the outline of the conducted experiment is mentioned. To answer the research questions and hypotheses, the results of the experiment, and input from students and teachers will be analysed in this chapter. This chapter will first describe a plan of analysis about what data has been gathered and how to use the data in order to answer the research questions and hypotheses. Furthermore the results will be split up into a qualitative and a quantitative analysis. The quantitative analysis will provide analysis to answer the hypothesis, the qualitative analysis will provide information about behavior and experiences from the students, teachers and researcher.

6.1. Plan of Analysis

In order to properly analyse the results, we formulated a plan for analysis, containing research questions, related variables and research units.

In the conducted experiment, a test and control group were identified, representing feedback and non-feedback teaching strategies. From here on, the test group will be called the feedback group, for simplicity and clarity. Within both groups, three tests were conducted to assess the performance of the students during the process. Besides the math tests, tests were conducted to assess the emotions and experiences of the students. Consequently, a list of variables could be composed.

Concerning our research questions and hypothesis as stated in chapter 1, and the claims based on the use cases as stated in section 3.4 important data variables are:

- Session results (Algorithm). How many times did NAO provide feedback, did students perform better in the second session?
- PANAS results (Affection scores). The results gathered from the PANAS tests before the session, after the first session and after the second session: PanasPre, PanasPost1, PanasPost2.
- Math test results. The results of the math tests done in class before the sessions, directly after the sessions and two weeks after the sessions: TestPre, TestPost1, TestPost2.
- Interaction Mechanism (Data gathered by NAO): Times the feet were touched, times the head was touched, number of exercises done, etc.
- The answers to the survey, especially questions about the motivational dance and the feedback given.
- Observations (Data gathered by the researcher): Mistakes NAO made, behavior of the student, etc.

An entire list of gathered variables can be found in appendix K.

The main purpose of the analysis is to examine the relationship between the given feedback by NAO and the performance of the student. As there are multiple test points and multiple groups, the Mixed ANOVA method is suitable to examine this relationship. Besides the given feedback by NAO, other factors may influence the math performance of the students, such as the perceived emotions, the quality of the feedback and the performance of NAO. These factors will be analysed separately.

As stated in chapter 1, our hypotheses: Feedback provided by NAO will improve the student math test results, and: Feedback increased the affection of the student, will be tested. Furthermore it is of interest to see if a low or negative PANAS result has had effect on the performance growth. We expect to see more feedback given in the group of students that scores low on the tests, and see less feedback in the second session compared to the first. During the analysis, the initial math grade is disregarded, because it also contains information about other math subjects than calculations.

Apart from the collected quantitative data, also the questionnaires, and experiences and observation of the researcher provided information. This will be analysed using qualitative methods.

6.2. Quantitative Analysis

Before testing the hypotheses, we take a look at the performance of NAO. How many times was feedback given? How many hearing mistakes did NAO make? For testing the hypotheses, we are using Mixed ANOVA in order to analyse the data. First, we test whether the feedback group actually did better on the tests than the control group, meaning the feedback would be effective. After that, we test whether the feedback group has more affection towards NAO than the control group.

6.2.1. Feedback Algorithm

NAO recognised a total of 15 mistakes in 143 exercises (10,5%) by the feedback group in the first session and 12 mistakes in 139 exercises (8,6%) by the feedback group in the second session. All classified mistakes (mistakes that could be detected by NAO as described in section 4.3) that were made by the students, were recognised by NAO (100% of the classified mistakes were recognised). The other mistakes that were made by students were unclassified, meaning NAO could not provide adapted feedback according to the answer. NAO was able to provide an explanation of the exercise 78 times in the first session, compared to 80 times in the second session. Remind that an explanation of the exercise is always given to the feedback group after a second wrong answer. This does not necessarily mean that the first wrong answer was a mistake NAO (should have) recognised.

15 times in the first session, adapted feedback was given to the student in the feedback group. After this feedback, 8 exercises were answered correctly. The feedback was effective 53% of the times given. In the second session, this effectiveness drops to 4 out of 12 (33%). We see that this drop takes effect mainly in the students who have a low grade in math. In the first session, 8 times feedback about lacing was given, 4 times about the operator, and three times about split and add. In the second session, 4 times feedback about lacing was given, 1 time about the operator and 7 times about split and add. In the first as well as the second session, we see that the lacing feedback was most effective, while the split and add feedback was least effective.

6.2.2. PANAS

We use mixed ANOVA to test whether the feedback group has a higher PANAS score than the control group, and thus has a more positive affection towards NAO. We do this by testing the hypothesis H0: The feedback group has a more positive affection towards NAO than the control group, against H1: The feedback group does not have a more positive affection towards NAO than the control group.

Descriptive Statistics

Group		Mean	Std. Deviation	N
PanasPre	Control Group	10,90	7,949	21
	Feedback Group	12,20	8,395	20
PanasPost1	Control Group	15,76	8,734	21
	Feedback Group	15,85	10,384	20
PanasPost2	Control Group	14,52	8,976	21
	Feedback Group	14,75	9,552	20

Table 6.1: Descriptive statistics of the PANAS scores.

Table 6.1 shows the mean and standard deviation descriptive statistics from the PANAS scores.

Overall the feedback group has higher scores than the control group, with an average score of 12,20, 15,85 and 14,75 compared to 10,90, 15,76 and 14,52 respectively.

In order to test the hypothesis H0, we first need to know whether sphericity is violated. To test this, we apply Mauchly's sphericity test. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi^2 = 7,797$, $p = 0,020$. Because of that, the degrees of freedom have been adjusted using Greenhouse-Geisser ($\epsilon = 0,844$).

As the sphericity was violated, we use the results with the adjusted degrees of freedom. A test of within-subjects effects showed that the main effect of the PANAS scores was significantly different over time. Main effect of the PANAS, $F(1,687\ ,\ 65,794) = 5,460$, $p = 0,009$.

A test of between-subjects effects shows that there are no significant differences between the feedback group and the control group $p = 0,822 > 0,05$. The test of within-subjects effects also shows that there is no significant interaction-effect between the PANAS scores and the groups. $F(1,687\ ,\ 65,794) = 0,123$, $p = 0,851$. This means that there is no difference in the PANAS scores for the two different groups.

If we look at contrasts between the different measuring points, we find that the first PANAS test was significantly lower than the second PANAS test ($p = 0,002$), but the second PANAS test was not significantly different compared to the last PANAS test ($p = 0,282$). The contrasts also show that there are no significant differences in the PANAS scores between both groups from the first to the second and from the second to the third PANAS test. The results can be seen in figure 6.1.

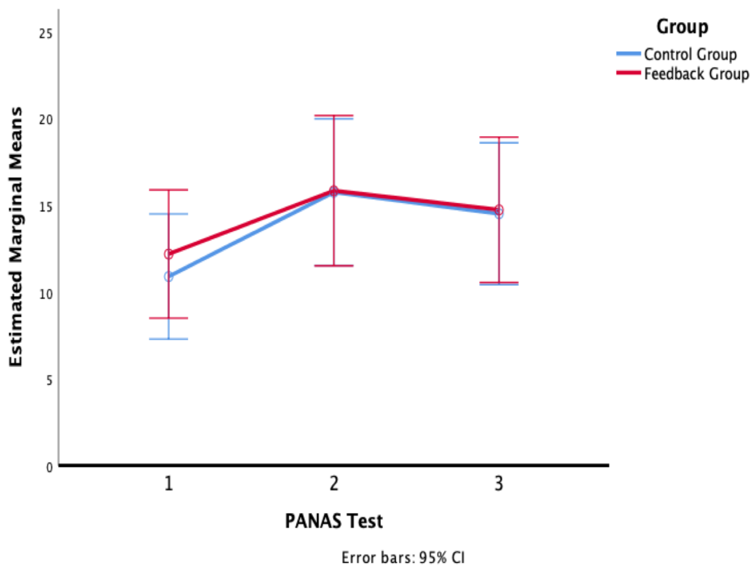


Figure 6.1: Estimated marginal means of the PANAS scores.

Correlations

		PanasPost1
correct_per1	Pearson Correlation	-0,082
	Sig. (2-tailed)	0,610
	N	41

Table 6.2: Correlation table of correct_per1 and PanasPost1.

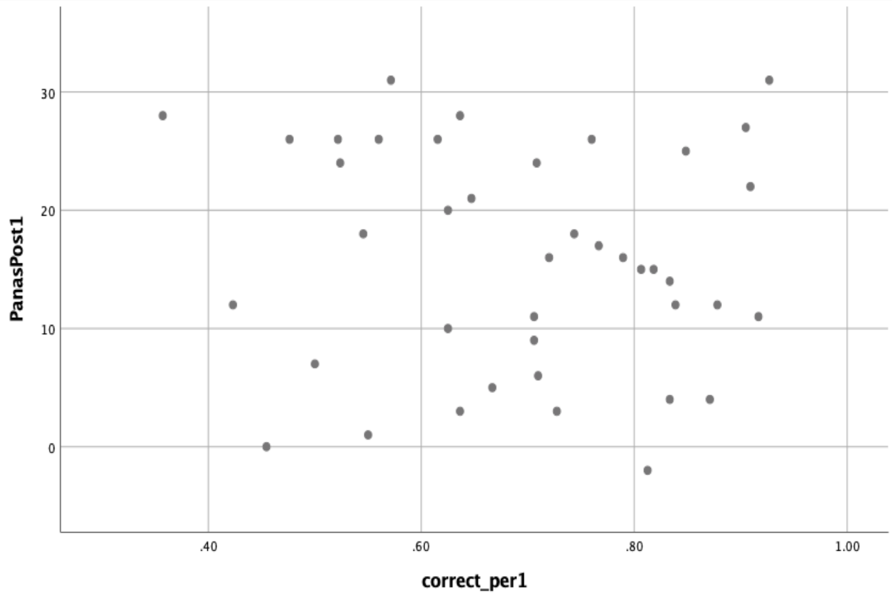


Figure 6.2: Scatterplot of PanasPost1 against correct_per1.

Correlations

		PanasPost2
correct_per2	Pearson Correlation	0,025
	Sig. (2-tailed)	0,878
	N	41

Table 6.3: Correlation table of correct_per2 and PanasPost2.

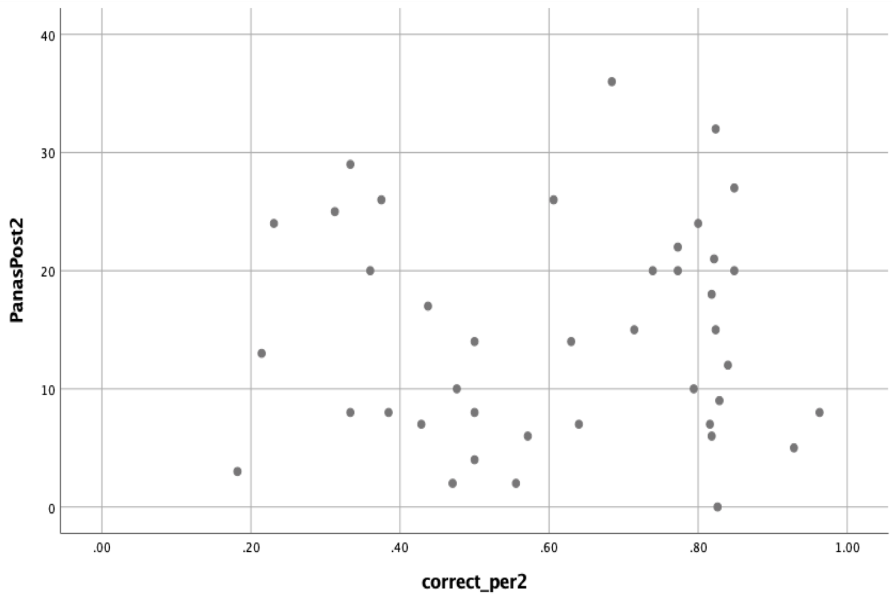


Figure 6.3: Scatterplot of PanasPost2 against correct_per2.

To see if the PANAS score has an effect on the answers given by the students during the math session with NAO, we try to correlate each PANAS test, with the session the student did. Because the PANAS results were not significantly different for both groups, we combine them. Correct_per is the ratio of correct answers (in one try), divided by the number of exercises done.

$$correct\_per = \frac{\sum total\_0}{\sum exercises}$$

In table 6.2 and figure 6.2, we can see the results of the correlation and the scatterplot belonging to it. PanasPost1 and correct_per1 correlate with a value of $-0,082$ where $p = 0,610$. This is not significant, therefore, we cannot find a correlation between the affection of a student towards NAO in the first session and the results during session 1. In table 6.3 and figure 6.3, we can see the results of the correlation and the scatterplot belonging to it. PanasPost2 and correct_per2 correlate with a value of $0,025$ where $p = 0,878$. This is not significant, therefore, we also cannot find a correlation between the affection of a student towards NAO in the second session and the results during session 2.

6.2.3. Math tests

We use Mixed ANOVA to test the hypothesis H0: The feedback group has a better math test score than the control group, against H1: The feedback group does not have a better math test score than the control group. In figures 6.4, 6.5 and 6.6 the results of the control group and feedback group can be found. Descriptive statistics can be found in table 6.4. The descriptive statistics as well as the bars seem to give little differences between the feedback group and the control group. The math tests have a range of [0, 20].

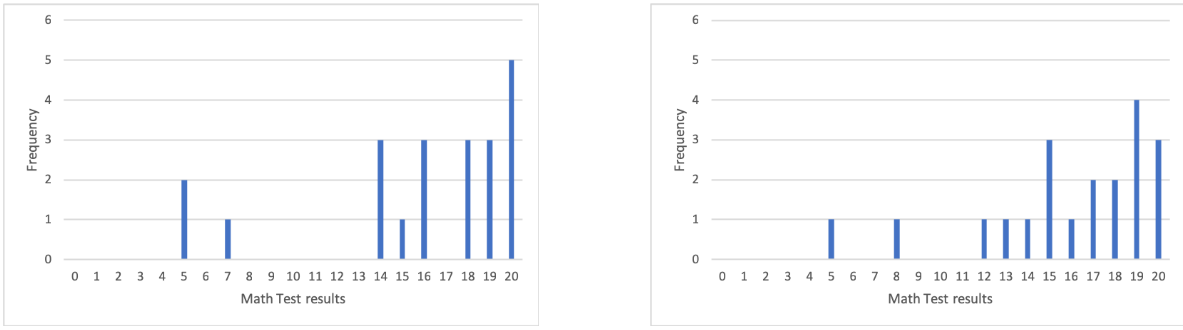


Figure 6.4: The results from the pre math tests. Left control group, right feedback group.

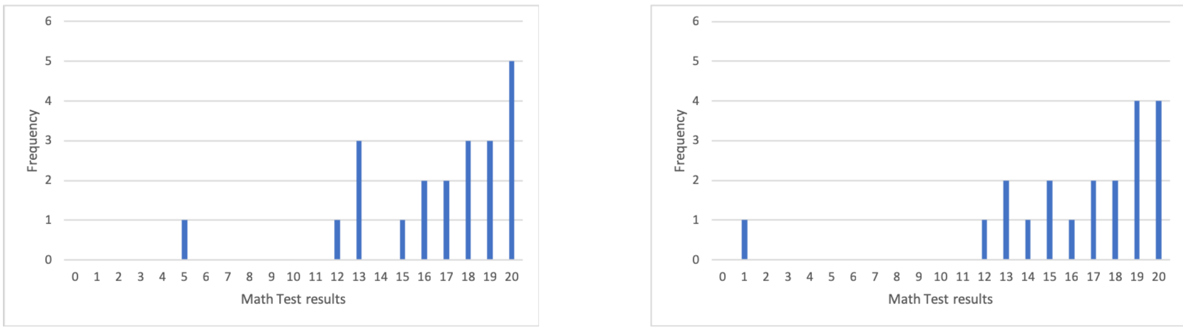


Figure 6.5: The results from the first post math tests. Left control group, right feedback group.

Descriptive Statistics

Group		Mean	Std. Deviation	N
TestPre	Control Group	15,86	4,767	21
	Feedback Group	15,95	4,045	20
TestPost1	Control Group	16,57	3,736	21
	Feedback Group	16,25	4,435	20
TestPost2	Control Group	17,00	2,846	21
	Feedback Group	15,35	5,264	20

Table 6.4: Descriptive statistics of the math test scores.

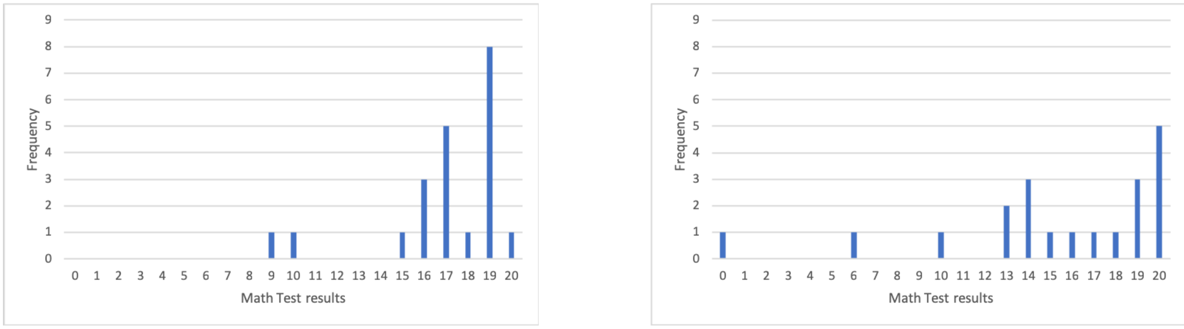


Figure 6.6: The results from the second post math tests. Left control group, right feedback group.

In order to test the hypothesis H0, we first need to know whether sphericity is violated. This is important to cancel out potential increase of Type I errors. To test this, we apply Mauchly's sphericity test. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2 = 2,471$, $p = 0,291$. As the sphericity was not violated, we use the results assuming sphericity. A test of within-subjects effects showed that the main effect of the math test scores was not significantly different over time. Main effect of the math tests, $F(2\ ,78) = 0,465$, $p = 0,630$.

A test of between-subjects effects shows that there are no significant differences between the feedback group and the control group $p = 0,598 > 0,05$. The test of within-subjects effects also shows that there is no significant interaction-effect between the math test scores and the groups. $F(2\ ,78) = 1,496$, $p = 0,230$. This means that there is no difference in the effect of the math test scores for the two different groups.

If we look at contrasts between the different measuring points, we see that the first math test was not significantly lower than the second math test ($p = 0,316$), and the second math test was also not significantly different compared to the last math test ($p = 0,630$). The contrasts also show that there are no significant differences in the math scores between both groups from the first to the second and from the second to the third math test. The results are displayed in figure 6.7, and a zoomed in version in 6.8.

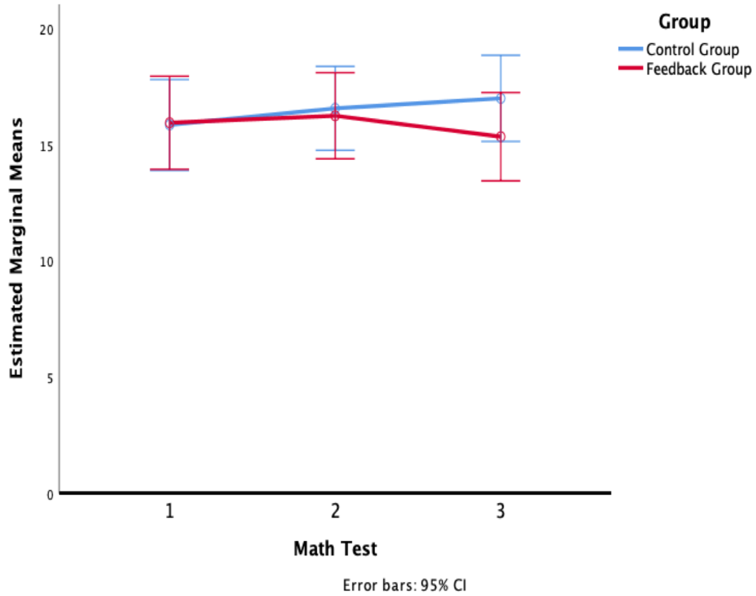


Figure 6.7: Estimated marginal means of the math scores.

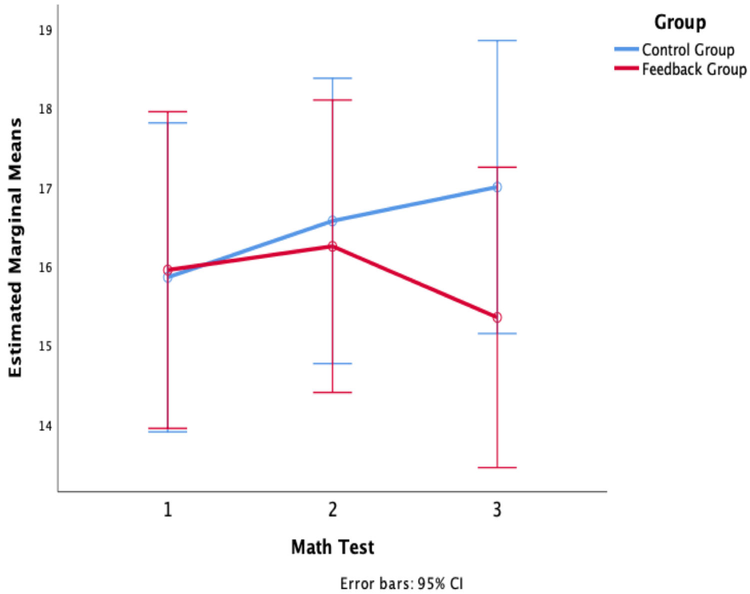


Figure 6.8: Zoomed in estimated marginal means of the math scores.

6.2.4. Interaction Mechanism

In the first session, NAO's feet were touched 2171 times, over a total of 1115 exercises. In the second session, NAO's feet were touched 1818 times, over a total of 935 exercises. On average, the feedback group touched the feet 2,1 times per exercise in the first session, and 2,2 times per exercise in the second session. The control group touched the feet 2,0 times per exercise in the first session, and 2,1 times per exercise in the second session. We use Paired Samples T-Test to see if we can find significant differences between the first and the second session. The feedback group did not touch the feet significantly more per exercise in the first session, compared to the second session $t(19) = -0,364$, $p = 0,72$. The control group did also not touch the feet significantly more per exercise in the first session, compared to the second session $t(20) = -0,707$, $p = 0,49$.

On average, students of the control group were too late giving an answer 2,81 times in the first session, compared to 0,52 times in the second session. Students of the feedback group were too late giving an answer 2,20 times in the first session, compared to 0,80 times in the second session. We use Paired Samples T-Test to see if we can find significant differences between the first and the second session. The feedback group was not significantly fewer times too late answering in the second session, compared to the first session $t(19) = 2,041$, $p = 0,055$. The control group was significantly fewer times too late answering in the second session, compared to the first session $t(20) = 2,503$, $p = 0,021$.

NAO made 137 mistakes in listening to the student, by misunderstanding a correct answer in the first session. On average, in the first session NAO made 3,4 mistakes per student in the first session, and 1,7 mistakes in the second session. We use Paired Samples T-Test, and find that NAO made significantly less mistakes in the second session, compared to the first session in the control group, $t(20) = 3,971$, $p = 0,001$. In the feedback group, NAO made 3,3 mistakes per student in the first session, and 1,9 mistakes per student in the second session. We use Paired Samples T-Test, and find that NAO made significantly less mistakes in the second session, compared to the first session in the feedback group. $t(19) = 2,547$, $p = 0,020$. In some cases, NAO was only able to understand half of the answer given by a student. In 18 cases in the first session, this led to a wrong answer for the feedback group, meaning the feedback group might have gotten feedback unnecessarily. In the second session, in 14 cases NAO understood only half the answer. This could be seen in the answers NAO interpreted, by getting only half of the answer (e.g. 72 is interpreted as 70 or 2).

Students of the control group touched NAO's head 0,26 times per exercise in the first session, and 0,50 times per exercise in the second session to request a repetition of the exercise. We use Paired Samples T-Test, and find that NAO's head was touched significantly more in the second session, compared to the first session in the control group. $t(20) = -4,387$, $p = 0,0002$. Students of the feedback group touched NAO's head 0,29 times per exercise in the first session, and 0,55 times per exercise in the second session to request a repetition of the exercise. We use Paired Samples T-Test, and find that NAO's head was touched significantly more in the second session, compared to the first session in the control group. $t(19) = -4,833$, $p = 0,0001$.

On average, students of the control group needed 42,3 seconds per exercise in the first session, compared to 54,0 seconds in the second session. We find a significant growth in the time needed using Paired Samples T-Test. $t(20) = -3,744$, $p = 0,001$. Students of the feedback group needed 55,8 seconds per exercise in the first session, compared to 72,0 seconds per exercise in the second session. We find a significant growth in the time needed using Paired Samples T-Test. $t(19) = -2,505$, $p = 0,021$. Students of the control group have done 30,4 exercises on average in the first session, compared to 24.5 exercises in the second session. We find a significant drop in the number of exercises done by the control group using Paired Samples T-Test. $t(20) = 4,353$, $p = 0,0003$. Students of the feedback group have done 23,9 exercises on average in the first session, compared to 21 exercises in the second session. We find no significant difference in the number of exercises done by the feedback group, using Paired Samples T-Test. $t(19) = 1,594$, $p = 0,127$. When we take a look at the differences between the groups, we find a significant difference between the feedback group and the control group using One-way ANOVA. The control group did significantly more exercises in the first session than the feedback group. $F(1,39) = 7,321$, $p = 0.010$. In the second session, there is no significant difference between the feedback group and the control group. $F(1,39) = 1,679$, $p = 0.203$.

6.2.5. Student Survey

The student survey contained questions about working with NAO, math and robots in common. The survey can be found in appendix G. The questions that contained a likert-scale were analysed using Paired Samples T-Test. No significant differences can be found in the feedback group for the questions in the first session compared to the second session. In the control group, we find a significant increase in the statement "Ik heb het idee dat Pixel mij iets heeft kunnen leren." (I think Pixel has taught me stuff), with an average of 3,81 out of 5 in the first session and 4,19 out of 5 in the second session. $t(20) = 2.359$, $p = 0,029$.

6.3. Qualitative Analysis

Besides the numbers in the previous section, it is of interest what the experiences of the children and teachers were. To do this, we mainly look at the behavior of the students during the sessions and their answers to the questionnaire, and the talks of the researcher with the teachers. The talks with teachers were informal and took place during lunch breaks, after school time and before the start of the day. These talks were mainly about NAO and the influence that teachers noticed.

6.3.1. Student Behavior

One of the most interesting parts to look at, is the behavior of the student during the sessions.

Some of the students were very shy, therefore looking down and talking softly. They had trouble communicating with the robot at first, but this increased in the second session already. Some students need some more time to feel at ease with NAO than others, they seem impressed, almost afraid of what it can do.

On the other hand, some students were extremely excited by what NAO was capable of. They started imitating it, touch it during the session or got excited by the rainbow eyes when they gave a correct answer. Some even tried out NAO's capabilities. They tried the gazing function, see how far it can follow the student.

Another very interesting turn of events was that although NAO tells the students to talk loud and clearly, students tend to start talking slower to pronounce certain numbers better when NAO does not understand the student. This led to NAO understanding only parts of the number. For example pronouncing thirtytwo as "thirty.......two", where NAO is only able to hear thirty. In Dutch, pronouncing 32 this way leads to *"twee...en...dertig"*, causing NAO to only understand *"twee"*, which is a wrong answer. This happened quite often, causing doubt and frustration for the students.

6.3.2. Student Input

Many students say they like NAO better than doing math on a computer, because a computer program has a timer most of the time. They are forced to calculate quickly, while some of the students may not be able to do so yet. NAO did not have a timer, therefore giving the student all the time (s)he needed per exercise to think and complete it successfully.

Both groups of students really liked the encouragement of a good answer. A couple of students even said ”Wow, rainbow eyes”, with a big smile, when they saw it for the first time, meaning they gave a correct answer. Only three students mentioned that they did not like the motivational dance at the end of the first session, where 5 students mentioned this at the end of the second session. Students who mentioned that they did not like the dance, said this was because of the dance they had to perform themselves. They mentioned that they did like an ending reward, however.

All students believed that doing math with NAO increased their calculation skills, although some of them mentioned that NAO did not teach them anything new. Only a couple of students did not like that NAO told the student what went wrong. They mentioned that it took NAO too long to explain the exercise, where they preferred to have a short explanation and continue with the next exercise.

After the sessions, students were asked for tips for NAO to be a better help the next time. Where most students provided the answer: ”NAO should listen better to my answer”, a difference was noticed by the researcher in answers from students who had a high math grade. These students were more likely to get frustrated by NAO’s slow talking, bad hearing, and the explanation took too long. This was also noticed during the session, where students who had a high math grade got frustrated more easily. The students expressed this by sighing, muttering about how slow NAO was, and during explanations muttering ”Yes, yes, yes”.

6.3.3. Teacher’s Experience

During and after the experiment period, the researcher had talks with the teachers of the fifth grade to see if they experience any differences or complications. All teachers at the school mentioned that the students were extremely hyped about working with NAO, and could not wait to do it again. The teachers also mentioned that NAO was subject to speech in class between the students, and even students from other grades got excited by their stories. Teachers from other grades were pitied that only the fifth grade could join a project like this, as they hoped that the level would adapt to other grades as well. All teachers at the school, young and old, hope that innovations like NAO will be adapted quickly in primary schools.

6.4. Conclusion

NAO was able to recognise 15 mistakes made by students in the first session and 12 mistakes made by students in the second session. 100% of the classified mistakes were recognised. In the first session students gave a correct answer after feedback was given 53% of the time, in the second session the number of correct answers after feedback was a bit lower, 33%. We used Mixed ANOVA to analyse the PANAS scores of the students, but did not find significant differences between the feedback and the control group. We try to find a relation between the PANAS scores and the number of correct answers given, by correlating these two variables. No correlation was found in as well the first as the second session. We used Mixed ANOVA to analyse the math test results as well, but did not find any significant differences between the feedback and the control group.

When we look at interaction, many significant differences can be seen. Fewer times, students of the control group were too late in giving an answer to NAO. Also, less hearing mistakes by NAO were made in both groups in the second session.

Students needed some time to feel at ease with the robot, but also were very curious to what NAO was able to do. Students had troubles using the speech recognition of NAO. Although the numbers might not reveal much, the experience of the students and the researcher reveals that speech recognition was a great blocker in many cases. Mostly students were very happy to work with NAO, liked the encouragement and although the analysis does not find a significant difference, students feel like they have learned from NAO.

7

Discussion

7.1. Algorithm

Only a small amount of the mistakes made by the students were classified as in section 3.6. The mistakes that were classified were all found and feedback was given accordingly. The 90% of the mistakes that were not recognised were either mistakes that were not classified in literature, or misunderstandings by NAO. More research to the thinking and learning process of students would help in classifying more mistakes. Perhaps more mistakes could have been detected when combinations of mistakes could have been detected by NAO. NAO provided less explanations of exercises in the second session compared to the first session. This can be caused by several reasons. The first reason is that the feedback was effective, and therefore less explanations might be needed. The other reason could be that less hearing mistakes were made, and therefore less students reached the explanation part. Which of the two reasons would be most presumable is hard to say, because the data lacks the opportunity to do so.

7.2. PANAS

The analysis did not provide significant differences for the affection of the students towards NAO between the feedback group and the control group. If we look at averages, the feedback group has a higher score than the control group, but no significant difference in growth or drop of the results are found between the two groups, meaning that we cannot conclude that the feedback provided by NAO had major effects on the students affection. We can therefore reject the hypothesis: Feedback increased the affection of the student.

We can also see that there is no correlation between the number of correct answers and the score on the PANAS test of the students. Quinn et al. mention that children who perform well in school may do so in part because they are happy, and performing well academically may make children happier [28], and Achor states that happiness increases performances in his seven principles of positive psychology [6]. The results of the correlation state that there is no difference between happy or unhappy students. Happy students can perform good and bad, as well as unhappy students. The novelty effect could be a reason for this. Nearly every student, good or bad at math, was excited to work with NAO. Many did not touch, see or work with a robot like NAO before. As both the control and the feedback group could work with NAO, no significant differences can be found between the groups. Although the first wow-effect was taken away by the introductory class, and students got to train with NAO before the sessions started, it still was a new environment to work in, which might have led to the fact that no differences in affection can be measured.

7.3. Math Tests

Statistical analysis did not provide any significant result about the math tests the students did. According to the results of the statistical analysis, no real improvement can be seen in the feedback group, compared to the control group. This means that the feedback by NAO seems not to have had any effect on the skills of the child. We can therefore reject the hypothesis: Feedback provided by NAO will improve the student's math test results. Differences are very small, so we can not conclude any improvement nor decrease of skills of the students in practising math. One of the reasons for this could be the ceiling effect. As most students already scored very high on the pretest, little improvements could be made by NAO's feedback. Another reason might be time. In four weeks time, it is hard to measure a major difference. Students need time to process and develop calculation skills, and thus their results. A third reason could be the group size. If more students would have entered the experiment, perhaps a bigger difference between the groups might have been found. One last reason might be that providing elaborated feedback as described in section 3.6 can ease the thinking process of a student. Following the model of Hattie and Timperley as described in section 2.5, NAO mainly provided task-related feedback, and some process-related feedback. A student who only receives a ”wrong” or ”correct” from NAO, needs to think for himself what mistake it made and where in the process, compared to a student who does not need to think about that and gets it presented by NAO. When a student needs to think about what went wrong, and where in the process this went wrong, it might lead to a better result, than actually needing to remember what NAO has taught the student.

7.4. Interaction Mechanism

The number of times students were able to give an answer to NAO, thus the amount of times the feet were touched, didn't change much for both groups in the first session compared to the second session. Interaction with NAO did increase over time. Less students were late in answering in the second session compared to the first session and NAO made less hearing mistakes, meaning the students got used to working with NAO quickly. On the other hand, NAO's speech recognition seems to be an issue during the sessions. Although the numbers are small, experiences of the researcher and students provide the feeling that NAO does not understand, or misunderstands the answer too often, which led to frustrations for the students.

We do however see an increase in the number of times the head is touched, over less exercises. Also, an increase in the time needed for an exercise in the second session is found in both the groups, compared to the first session. This could indicate that students needed more time, because the exercises came from categories they made most mistakes in during the first session. Because of that, an increase in the number of times the head was touched can be clarified.

When we look at the number of exercises done, we find that significantly less exercises were done in the control group in the second session, compared to the first session, but there was a small decrease in the feedback group as well. Reason might be that the control group had less interference of NAO, because the simple feedback took less time. When we look at the differences between the groups, the control group did more exercises in the first session, but no significant differences were found in the second session. None of the students had a clean sheet, meaning none of the students did not make any mistake. That indicates that the exercises were not too easy for the students to solve. When we combine all the significant differences above, we can indicate that the feedback group remained more stable between sessions, compared to the control group. Interference of NAO by providing feedback during exercises could be the reason for that.

Overall, the interaction mechanism was robust. At all times, students were able to successfully complete an exercise, and continue to the next one, without interference of the researcher.

7.5. Survey and Observational Data

The answers to the Likert-scale questions on the survey, were bound to the ceiling effect as well. On a Likert-scale ranged from 1 to 5, the averages were all 3.8 or higher, with small differences between the first and the second session.

When we take a look at the differences between the qualitative and quantitative analysis, some contradictions arise. Statistical analysis showed that there was no significant growth in the skills of a student, but all students mentioned that they felt that they did learn from the session with NAO. Extraordinarily, a significant difference in the control group was found according to the statement about how much Pixel had taught the student. This could indicate that students without feedback feel that practise makes perfect, and by practising the categories they made mistakes in, they have a feeling that NAO has taught them something. The more a student practises his/her math skills, the better he or she gets at it.

When students were asked for tips in improving NAO, there was a clear difference between students who scored high and students who scored low. Almost all students who scored high on math tests thought that the explanation of NAO took too long, NAO spoke too slowly or processing an answer took too much time. They were eager to quickly answer the next question. They tended to become frustrated more easily than students who didn't score that high on math tests. They expressed their frustration by sighing or muttering "yes, yes, yes" as if they wanted NAO to speed up and continue. This can be easily explained by the fact that students who score high on math tests, might experience the explanations of NAO as obsolete and rather continue and finish the test, as they knew the answers already.

Contrary to students who score high on math tests, students who encounter difficulties in calculation strategies might profit more from the explanation NAO gives and the pace at which this happens. Therefore, the results of the questionnaire show that these students were more excited about the explanations of NAO. They clearly seem interested in hearing the entire solution, as they can benefit and improve their math skills.

7.6. Limitations

This research was limited by two main factors: time, and speech. In a small amount of time (the experiment was conducted for 2 weeks per primary school), it is hard to find a significant increase in the results. NAO only has been able to provide feedback and help the students for 2 sessions of 20 minutes, while a teacher is continuously helping students to grow. The influence of a teacher is way bigger than NAO's, especially in a small period of time. Also, a teacher is able to adapt quickly according to the level of the student. The other limitation is the speech recognition of NAO. Although students have said that they preferred talking to NAO than typing on a tablet of computer, common speech recognition is not able to fulfil a complete session, without misunderstanding or not understanding the student yet. The last limitation is more a lack of knowledge rather than a limitation. A lack of knowledge about mistakes and the thinking/solving process of a student was a problem as well. Only a small percentage of the exercises, NAO was able to give elaborated feedback. And although a decrease in feedback can be seen in the second session, I think more mistakes should be able to be handled by NAO in order to give better elaborated feedback.

8 Conclusion

In chapter 1, research questions and hypothesis were stated. The first research question: ”How can a humanoid robot provide feedback after a mistake was made by a primary school student in a basic calculation exercise with natural numbers up until 100?” is answered in chapter 2. In 4th and 5th grade, students use and repeat various simple strategies and calculation strategies. The generally most frequently used strategies are lacing and splitting. Therefore, NAO focused on these two methods. NAO also aimed to identify commonly made mistakes such as switching units in bigger number, or confusion about symbols or strategies. An important element of this research is feedback, which concerns information about how we perform in efforts to reach goals. It requires information about the goal, the followed track and the next steps. Feedback seems to be most effective when it is self-regulated and intended for process. Students need tangible and timely feedback, not too early (to prevent overwhelming the student) and not too late. NAO therefore first gave feedback about the track, and after the second try feedback on the entire process. All of this led to the design and implementation as stated in chapters 3 and 4.

The second research question: ”How will the interaction mechanism of NAO be shaped, in order to communicate with the student in the tutoring session?” is also answered in chapters 3 and 4. Two scenarios were formulated, an introductory class to let the children interact and touch NAO, and a math tutoring session, as was used during the experiment. NAO behaved human-like, acting like a tutor to help the students in their matter. Interaction was specified as NAO only using speech and movements. Feedback was specified by mentioning the mistake the student made, and after a second try by explaining how to reach a correct answer. Rewards were implemented to keep the student motivated, by using speech and LEDs in the eyes of NAO. After the session, a motivational dance act was performed.

The third research question: ”How did feedback and the robot influence the performance, affection and interaction of the student in practising math?”, together with the hypotheses: ”Feedback provided by NAO will improve the student math test results.” and ”Feedback improved the affection of the student towards NAO.” are answered in chapter 6, and were discussed in chapter 7. We can conclude that the algorithm for detecting mistakes and giving feedback accordingly worked as implemented, but might have done a better job if combinations of mistakes would have been recognised as well. Besides that, improvements could be made to the algorithm if we had more knowledge about the learning and thinking process of the student. No differences between the feedback group and the control group were found in the PANAS scores, meaning the affection of the student towards NAO was not affected by the feedback. Every student had the possibility to work with NAO, in a new environment, which most likely is the reason for this result. Also, no differences could be found in the results of the math tests. Students from the feedback group as well as the control group did not perform significantly different over time. The ceiling effect of the initial test scores could be a reason for that, or the fact that feedback as provided by NAO was unsuccessful for the students. Both hypotheses could therefore be rejected.

The interaction mechanism was robust during the sessions. At all times, the students were able to continue to the next exercise, without interference of the researcher. Interaction with NAO improved a lot between the first and the second session. Less students were late in answering in the second session compared to the first session and NAO made less hearing mistakes, meaning the students got used to working with NAO quickly, but speech recognition seemed to be a limitation in both sessions, as was experienced by the students and researcher. NAO's interference by providing feedback might have caused that we found a difference between the control group and the feedback group when it comes to interaction. When NAO had less to tell the students, students were able to do more exercises, and proceed quicker.

Contradictions were found when we compared the qualitative analysis with the quantitative analysis. Statistical analysis showed that there was no significant growth in the skills of a student, but all students mentioned that they felt that they did learn from the session with NAO. We also find a difference between students who scored a high math grade, compared to students who did not. Students who scored high more often thought the explanation of NAO took too long, NAO spoke too slowly or processing an answer took too much time. They might experience the explanations of NAO as obsolete and rather continue and finish the test, as they knew the answers already, where the students who scored a low grade clearly seem interested in hearing the entire solution, as they can benefit and improve their math skills.

This research was limited by several factors: time, speech, and lack of educational knowledge. In a small amount of time (the experiment was conducted for 2 weeks per primary school), it is hard to find a significant increase in the results. Although students have said that they preferred talking to NAO than typing on a tablet of computer, common speech recognition is not able to fulfil a complete session, without misunderstanding or not understanding the student yet. Besides that, a lack of knowledge about mistakes and the thinking/solving process of a student was a problem as well.

To conclude, results did not show any difference between the group who was provided elaborated feedback by NAO and the group who was provided simple feedback. The limited time to do this research and small groups of students might have led to this result.

Students as well as teachers were very happy and excited to work with a robot like NAO, therefore I believe that the experiment was a success. I would suggest to adapt more direct feedback in math exercises for students. However, we will need to increase the level of NAO's adaptiveness in order to see more effect from its feedback.

8.1. Future Work

My research had a big potential, but was very limited by time. As said before, students as well as teachers were very happy and excited to work with a robot like NAO. However, if you want to measure improvements in study skills, four weeks is very little time. Influence of NAO on long-term memory is hard to be seen in four weeks. For further research, I would suggest to provide feedback by a humanoid robot for about a year, doing measurements at the beginning, halfway and at the end of the school year. If each student gets to work with NAO once a week, I think a clearer difference in growth should be seen, because you maximise the influence of NAO with respect to the teacher. Also make sure that the students are at ease with the robot, so that a constant line can be seen in the interaction mechanism.

Besides that, combining mistakes could lead to an increase in the feedback given, which might lead to positive results as well. Also, if NAO were able to understand more than the plain answer, NAO would in future maybe be able to find mistakes in the thinking process. That would also mean that students would need to think out loud, and NAO would need to understand a lot more than just plain math.

For a robot like NAO to really help out in schools, more adaptiveness should be applied. The level of exercises could be adapted to the knowledge of the child, the feedback given could be adapted. Once an exercise has been explained in a category, try to let the student think about what went wrong, and provide only part of the explanation instead of the entire story, to speed up the process. Perhaps ask in between parts of the explanation whether the student remembers now, and wants to continue or do the exercise again. As was mentioned earlier, a difference within the groups might be found as well. Conducting research where groups are created according to their level of math, providing the same feedback in both groups might show that NAO has more influence on students who have a hard time practising math.

A

Training script

NAO:

"Hallo, ik ben Pixel. Ik ben een rekenrobot. Ik heb wat problemen met horen, ik versta niet altijd goed wat je zegt. Om je goed te kunnen verstaan moet je hard en duidelijk praten. Dit gaan we nu even oefenen. Ik ga 1 voor 1, 5 getallen opnoemen. Als jij dan dit getal herhaalt, dan zeg ik wat ik gehoord heb. Je kan beginnen met praten, zodra mijn ogen blauw kleuren.
Het eerste getal is: 37
Het volgende getal is: 98
Het volgende getal is: 7
Het volgende getal is: 11
Het volgende getal is: 42"

If NAO didn't hear the student:
"Ik heb je niet gehoord. wil je het nog een keer zeggen?"
When the student answers:
"Volgens mij zei je [number]"
If the number is not the same as NAO said:
"Dat is niet wat ik zei. Probeer nog eens luid en duidelijk te praten."

After all numbers have been repeated correctly, start tickle game:
"Dat waren alle getalletjes! We gaan nu oefenen met het aanraken van mijn lichaam. Wanneer je mijn voeten of hoofd aanraakt, kietelt dat en moet ik heel hard lachen. Zullen we eens kijken of je mij aan het lachen kan krijgen?" "Ga je gang, probeer het maar!"
Tickle-game for one minute

"Bedankt voor het oefenen! Nu is het tijd om samen te gaan rekenen. We oefenen even 3 makkelijke rekensommetjes, om te kijken of het goed gaat. Denk eraan dat je alleen tegen mij kan praten als mijn ogen blauw zijn.
Hoeveel is: 1 erbij 1?
Hoeveel is: 1 erbij 2?
Hoeveel is: 2 erbij 2?"

In case of a correct answer:
"Heel goed!" In case of a wrong answer:
"Dat is helaas niet goed."
After the three exercises:
"Dit was de training. Dankjewel!"

B
Session script

If it is the first session:
NAO: ”Welkom! Mijn naam is Pixel. Wat is jouw naam?.”
Student: [student name].
NAO: ”Hallo! Leuk dat je met mijn sommetjes wil maken!
Vorig jaar heb je geleerd hoe je erbij en eraf sommen moet maken die tot 100 gaan.
Wij gaan 20 minuten lang rekensommetjes oefenen, die jij uit je hoofd moet proberen uit te rekenen.
Gaat dit goed, dan leer ik jou een dansje.
Maak je een foutje, dan probeer ik je uit te leggen, hoe het wel moet.
Vorig jaar heb je veel geoefend met het rekenen, vooral met erbij en eraf sommetjes die tot 100 gaan.
Je hebt geleerd hoe je over het tiental moet heenrekenen, zoals bijvoorbeeld bij 18 erbij 3, en wat er moet gebeuren als je 12 eraf 4 moet uitrekenen.
Vandaag gaan we verder met oefenen van deze sommetjes.

If it is the second session:
NAO: ”Welkom terug! Leuk dat je er weer bent!
We gaan weer verder met sommetjes oefenen, waar je de vorige keer nog wat moeite mee had. Ik zal je even uitleggen wat we gaan doen.

In both sessions:
Als je de som niet goed verstaan hebt, of je wil hem tijdens het rekenen nog een keer horen, druk dan zachtjes op mijn hoofd.
Als je denkt dat je het antwoord op de som weet, druk dan tegen een van mijn voeten aan, dan zal ik luisteren naar je antwoord.
Denk eraan dat je alleen tegen mij kan praten als mijn ogen blauw kleuren.”

”De eerste/volgende rekensom is: Hoeveel is [rekensom]?”
Student: [answer]
In case of a wrong answer:

- ”Dat is helaas niet goed. Ik denk dat je een tussenstapje vergeten bent. Probeer het nog eens een keer.”
- ”Jammer, bijna goed! Ik denk dat je een fout gemaakt hebt bij het splitsen. Hou er rekeneing mee dat je altijd het eerste getal, eraf, het tweede getal moet doen. Als dat niet kan, moet je gaan lenen. Probeer het nog eens een keer.”
- ”Je hebt het bijna goed. Ik denk dat je een foutje gemaakt hebt bij het verbeelden van je getallen. [getal] schrijf je als [getal1] en [getal2]. Probeer het nog eens een keer.”
- ”Helaas, niet goed. Ik vond deze ook heel lastig! Probeer het nog eens een keer.”

In case of a correct answer:

- ”Goed gedaan! Dat is inderdaad het goede antwoord!”
- ”Alweer goed! Je bent goed bezig!”
- ”Je bent een kanjer! Dat is het goede antwoord!”
- ”Ja, helemaal goed! Op naar de volgende!”

When NAO is going to explain how to get to the correct answer:
”Dat is helaas niet het goede antwoord. Ik zal je uitleggen hoe we tot het goede antwoord komen.”

- Addition, units>10: ”We tellen eerst de tientallen van het tweede getal, bij het eerste getal op. [step]. We kijken of de eenheden bij elkaar opgeteld meer dan 10 zijn. [step] is meer dan 10. We tellen door tot een heel tiental. [step]. Daarna tellen we de rest erbij op. [step]. Het antwoord is dus [answer].”
- Addition, units<10: ”We tellen eerst de tientallen van het tweede getal, bij het eerste getal op. [step]. We kijken of de eenheden bij elkaar opgeteld meer dan 10 zijn. In dit geval, is dat niet zo. We kunnen dus de eenheden gemakkelijk bij elkaar optellen. [step]. Het antwoord is dus [answer].”
- Subtraction, units<0: ”We trekken eerst de tientallen van het tweede getal af van het eerste getal. [step]. We trekken de eenheden van elkaar af. In dit geval kunnen we niet zomaar de eenheden van elkaar aftrekken. [wrong step]. Dit kan niet. We moeten dus even een tiental lenen. Dit trekken we van [number] af. Nu kunnen we de som wel uitrekenen. [step]. Nu tellen we de twee stappen bij elkaar op, om tot het goede antwoord te komen. [step]. Het antwoord is dus [answer].”
- Subtraction, units>0: ””We trekken eerst de tientallen van het tweede getal af van het eerste getal. [step]. We trekken de eenheden van elkaar af. [step]. Het antwoord is dus [answer].”

”De tijd zit erop. Dat was leuk toch?
Ik zie je snel weer, dan gaan we weer verder oefenen.
Dankjewel voor al het harde werk. Zullen we nog een keer het dansje doen?
Dan ga ik daarna even rusten.”

C

Filelist

C.1. Training package

The training package consists of the following files:

- initialize.py
- training.py
- touch.py

The initialize.py file contains all variables and initial settings for NAO to be used. In order to use NAO's built-in functions, proxys are used to connect to the specific part in the NAOqi API. The proxys used are:

- ALTextToSpeech - Language: Dutch, speed: 90%
- ALAnimatedSpeech
- ALSpeechRecognition - Language: Dutch, vocabulary: numbers from 0 - 100, confidence set to 0.45
- ALMemory
- ALMotion
- ALTracker - Target: facetracking
- ALRobotPosture
- ALLeds
- ALBroker - connected to Touch events

The touch.py is an event listener. Whenever NAO is touched, an event is raised. This is used for a so-called "Tickle game". A game in which the children have to touch NAO, after which it will start laughing.

The training.py is the executable file which combines all files together. It contains methods to explain what it is about to do with the student, a method to activate speech recognition and listen to the answer of the student and some methods to start the tickle game.

C.2. Feedback Package

The feedback package consists of the following files:

- converse.py
- exercise.py
- initialize.py
- math.py
- reward.py
- session1.py
- student.py
- touch.py
- exercises.json

The converse.py file contains all simple conversation methods. It contains all pure textual phrases that NAO will say to the student.

The exercise.py file contains all methods concerning the exercises for the students. Methods to get all exercises, answers, current exercise and to check if the answer of the student is correct.

The initialize.py file contains all variables and initial settings for NAO to be used. In order to use NAO's built-in functions, proxies are used to connect to the specific part in the NAOqi API. The proxys used are:

- ALTextToSpeech - Language: Dutch, speed: 90%
- ALAnimatedSpeech
- ALSpeechRecognition - Language: Dutch, vocabulary: numbers from 0 - 100, confidence set to 0.45
- ALMemory
- ALMotion
- ALTracker - Target: facetracking
- ALRobotPosture
- ALBroker - connected to Touch events

The math.py file contains all math algorithms and possible mistakes made. It contains methods for mistakes made in split and add, lacing, and the possibility of switching numbers. Besides that, an algorithm of lacing is created to explain the student how to get to a correct answer.

The reward.py file contains all possible rewards NAO provides the student. The dance steps are created, as well as the rainbow eyes and the cheering.

The session1.py is the executable file, which imports all other files. It contains methods to provide an exercise, and a main method to work with the student.

The student.py file contains methods with speech recognition. It contains a method to hear and translate the answer of the student, and warns it when it takes a long time.

The touch.py is an event listener. Whenever NAO is touched, an event is raised. This is used for repeating and answering an exercise.

C.3. Non-feedback Package

The NAO Package without feedback contains the same methods and files as the feedback package, except for math.py. The students will get rewards, and will be told whether the answer given was correct or not. The explanation and elaborated feedback are left out.

D

Pretest

7 + 6 =	13 Over het tiental heen 7 plus 3 en dan nog 3 erbij	12 - 5 =	7 Door het tiental heen 12 - 2 en dan nog 3 eraf
8 + 8 =	16 Over het tiental heen 8 plus 2 en dan nog 6 erbij	16 - 9 =	7 Door het tiental heen 16 - 6 en dan nog 3 eraf
19 + 5 =	24 Over het tiental heen 19 plus 1 en dan nog 4 erbij	29 - 6 =	23 Tiental wegdenken en er later weer bij plaatsen 9 - 3 = 6 dus 29 - 3 = 26
17 + 8 =	25 Over het tiental heen 17 plus 3 en dan nog 5 erbij	53 - 2 =	51 Tiental wegdenken en er later weer bij plaatsen 3 - 2 = 1 dus 53 - 2 = 51
37 + 31 =	68 Tientallen plus eenheden erbij 30 + 30 = 60 1 + 7 = 8 60 + 8 = 68	38 - 17 =	21 Tientallen min tientallen en eenheden 38 - 10 = 28 28 - 7 = 21
27 + 42 =	69 Tientallen plus eenheden erbij 20 + 40 = 60 2 + 7 = 9 60 + 9 = 69	53 - 21 =	32 Tientallen min tientallen en eenheden 53 - 20 = 33 33 - 1 = 32
67 + 14 =	81 Tientallen plus eenheden erbij, over het tiental heen 67 + 3 = 70 70 + 1 = 71 71 + 10 = 81	77 - 17 =	60 Tientallen min tientallen en eenheden 77 - 10 = 67 67 - 7 = 60
38 + 27 =	65 Tientallen plus eenheden erbij, over het tiental heen 38 + 2 = 40 40 + 5 = 45 45 + 20 = 65	46 - 18 =	28 Tientallen en eenheden, door het tiental heen 46 - 6 = 40 40 - 2 = 38 38 - 10 = 28
24 + 76 =	100 Beginnen met het kleinste getal, tientallen plus eenheden erbij Omdraaien, 76 + 24 = 76 + 4 = 80 80 + 20 = 100	71 - 29 =	42 Tientallen en eenheden, door het tiental heen 71 - 1 = 70 70 - 8 = 62 62 - 20 = 42
13 + 57 =	70 Beginnen met het kleinste getal, tientallen plus eenheden erbij Omdraaien, 57 + 13 = 57 + 3 = 60 60 + 10 = 70	93 - 27 =	66 Tientallen en eenheden, door het tiental heen 93 - 3 = 90 90 - 4 = 86 86 - 20 = 66

E

Posttest

8 + 6 =	14 Over het tiental heen 8 plus 2 en dan nog 4 erbij	11 - 5 =	6 Door het tiental heen 11 - 1 en dan nog 4 eraf
8 + 9 =	17 Over het tiental heen 8 plus 2 en dan nog 7 erbij	15 - 9 =	6 Door het tiental heen 15 - 5 en dan nog 4 eraf
17 + 5 =	22 Over het tiental heen 16 plus 4 en dan nog 2 erbij	39 - 4 =	35 Tiental wegdenken en er later weer bij plaatsen 9 - 4 = 5 dus 39 - 4 = 35
18 + 6 =	24 Over het tiental heen 18 plus 2 en dan nog 4 erbij	64 - 3 =	61 Tiental wegdenken en er later weer bij plaatsen 4 - 3 = 1 dus 64 - 3 = 61
37 + 21 =	58 Tientallen plus eenheden erbij 30 + 20 = 50 7 + 1 = 8 50 + 8 = 58	58 - 32 =	26 Tientallen min tientallen en eenheden 58 - 30 = 28 28 - 2 = 26
42 + 37 =	79 Tientallen plus eenheden erbij 40 + 30 = 70 2 + 7 = 9 70 + 9 = 79	33 - 11 =	22 Tientallen min tientallen en eenheden 33 - 10 = 23 23 - 1 = 22
57 + 14 =	71 Tientallen plus eenheden erbij, over het tiental heen 57 + 3 = 60 60 + 1 = 61 61 + 10 = 71	66 - 36 =	30 Tientallen min tientallen en eenheden 66 - 30 = 36 36 - 6 = 30
47 + 37 =	84 Tientallen plus eenheden erbij, over het tiental heen 47 + 3 = 50 50 + 4 = 54 54 + 30 = 84	57 - 19 =	38 Tientallen en eenheden, door het tiental heen 57 - 7 = 50 50 - 2 = 48 48 - 10 = 38
26 + 34 =	60 Beginnen met het kleinste getal, tientallen plus eenheden erbij Omdraaien, 34 + 26 = 34 + 6 = 40 40 + 20 = 60	86 - 18 =	68 Tientallen en eenheden, door het tiental heen 86 - 6 = 80 80 - 2 = 78 78 - 10 = 68
23 + 47 =	70 Beginnen met het kleinste getal, tientallen plus eenheden erbij Omdraaien, 47 + 23 = 47 + 3 = 50 50 + 20 = 70	83 - 27 =	56 Tientallen en eenheden, door het tiental heen 83 - 3 = 80 80 - 4 = 76 76 - 20 = 56

F

Posttest2

8 + 5 =	13 Over het tiental heen 8 plus 2 en dan nog 3 erbij	11 - 7 =	4 Door het tiental heen 11 - 1 en dan nog 6 eraf
9 + 9 =	18 Over het tiental heen 9 plus 1 en dan nog 8 erbij	15 - 6 =	9 Door het tiental heen 15 - 5 en dan nog 1 eraf
17 + 4 =	21 Over het tiental heen 17 plus 3 en dan nog 1 erbij	39 - 6 =	33 Tiental wegdenken en er later weer bij plaatsen 9 - 6 = 3 dus 39 - 6 = 33
18 + 5 =	23 Over het tiental heen 18 plus 2 en dan nog 3 erbij	64 - 2 =	62 Tiental wegdenken en er later weer bij plaatsen 4 - 2 = 2 dus 64 - 2 = 62
57 + 21 =	78 Tientallen plus eenheden erbij 50 + 20 = 70 7 + 1 = 8 70 + 8 = 78	58 - 34 =	24 Tientallen min tientallen en eenheden 58 - 30 = 28 28 - 4 = 24
52 + 37 =	89 Tientallen plus eenheden erbij 50 + 30 = 80 2 + 7 = 9 80 + 9 = 89	33 - 12 =	21 Tientallen min tientallen en eenheden 33 - 10 = 23 23 - 2 = 21
57 + 24 =	81 Tientallen plus eenheden erbij, over het tiental heen 57 + 3 = 60 60 + 1 = 61 61 + 20 = 81	55 - 23 =	32 Tientallen min tientallen en eenheden 55 - 20 = 35 35 - 3 = 32
37 + 17 =	54 Tientallen plus eenheden erbij, over het tiental heen 37 + 3 = 40 40 + 4 = 44 44 + 10 = 54	57 - 39 =	18 Tientallen en eenheden, door het tiental heen 57 - 7 = 50 50 - 2 = 48 48 - 30 = 18
16 + 34 =	50 Beginnen met het kleinste getal, tientallen plus eenheden erbij Omdraaien, 34 + 16 = 34 + 6 = 40 40 + 10 = 50	86 - 19 =	67 Tientallen en eenheden, door het tiental heen 86 - 6 = 80 80 - 3 = 77 77 - 10 = 67
23 + 67 =	90 Beginnen met het kleinste getal, tientallen plus eenheden erbij Omdraaien, 67 + 23 = 67 + 3 = 70 70 + 20 = 90	83 - 47 =	46 Tientallen en eenheden, door het tiental heen 83 - 3 = 80 80 - 4 = 86 86 - 40 = 46

G

Survey

1. Heb je wel eens eerder een robot gezien?
2. Doen jullie hier op school veel met techniek?
3. Doe jij zelf veel met techniek?
4. Rekenen met Pixel vond ik leuk.
   (a) Vind je het leuker om met Pixel te oefenen dan uit het boek?
   (b) Vind je het leuker om met Pixel te oefenen dan met de computer?
5. Ik heb het idee dat Pixel mij iets heeft kunnen leren.
6. Door het oefenen met Pixel ben ik beter geworden in rekensommetjes maken.
7. Als ik weer met Pixel zou mogen rekenen, zou ik dat doen.
8. Dat Pixel mij een dansje leert, zorgde ervoor dat ik tot het einde wilde doorgaan.
9. Dat Pixel mij aanmoedigde bij een goed antwoord vond ik fijn.
10. -TESTGROUP ONLY- Dat Pixel mij vertelde wat ik fout deed vond ik fijn.
11. -TESTGROUP ONLY- Dat Pixel mij vertelde hoe ik de som moest uitrekenen vond ik fijn.

H

PANAS

Leerling nummer:
Pre / Post1 / Post2

	1 **Heel weinig of helemaal niet**	**2** **Een beetje**	**3** **Matig**	**4** **Veel**	**5** **Heel veel**
1. Geïnteresseerd	1	2	3	4	5
2. Bedroefd	1	2	3	4	5
3. Opgewekt	1	2	3	4	5
4. Sip	1	2	3	4	5
5. Sterk	1	2	3	4	5
6. Schuldig	1	2	3	4	5
7. Angstig	1	2	3	4	5
8. Vijandig	1	2	3	4	5
9. Enthousiast	1	2	3	4	5
10. Zelfverzekerd	1	2	3	4	5
11. Vlug geïrriteerd	1	2	3	4	5
12. Alert	1	2	3	4	5
13. Beschaamd	1	2	3	4	5
14. Vol inspiratie	1	2	3	4	5
15. Gespannen	1	2	3	4	5
16. Vastberaden	1	2	3	4	5
17. Aandachtig	1	2	3	4	5
18. Zenuwachtig	1	2	3	4	5
19. Energiek	1	2	3	4	5
20. Bang	1	2	3	4	5

Researcher Form

Student number:
Session number:

Aantal keer getraind:
Training score (1-3):

Aantal antwoorden goed door leerling, fout gerekend door Pixel:

Aantal antwoorden fout door leerling, goed gerekend door Pixel:

Te laat met antwoord geven:

Vergeten voet in te drukken bij antwoord geven:

Aantal keer voet gedrukt:

aantal keer hoofd gedrukt:

Aantal keer om hulp gevraagd:

Leerling gaat staan:

Leerling doet bewegingen robot na:

Leerling is geërgerd aan niet verstaan robot:

Leerling is geërgerd aan langzame tempo robot:

Andere gedragsobservaties:

Protocol:

- De training score wordt gegeven op een schaal van 1-3.
  - 1: Training ging voorspoedig, weinig tot geen fouten of hulp nodig gehad.
  - 2: Training ging matig, leerling maakte wat fouten, maar herstelde zichzelf of met wat hulp. Vertrouwen genoeg om zonder problemen de sessie door te komen.
  - 3: Training ging slecht, leerling maakte veel fouten en kon zonder hulp niet door. Geen vertrouwen in een goed verloop van de sessie. Training moet een tweede keer gedaan worden om vertrouwen te krijgen op een goed verloop.
- Observaties worden geturfd.
- In principe zitten leerlingen voor de robot. Echter komt het voor dat leerlingen gaan staan om op ooghoogte te komen met de robot. Dit wordt genoteerd als “Leerling gaat staan”.
- Indien de robot een aantal keer de leerling niet verstaat, kan het voorkomen dat een leerling zich gaat ergeren of ongeduldig wordt. Dit wordt genoteerd als “Leerling is geërgerd aan niet verstaan robot”.
- Indien de robot bij een aantal foute antwoorden de leerling gaat uitleggen hoe het wel moet, kan dit lang duren. Een leerling kan zich hieraan gaan ergeren of ongeduldig worden. Dit wordt genoteerd als “Leerling is geërgerd aan niet verstaan robot”.
- Andere gedragsobservaties worden onderaan genoteerd.

J

Experiment Protocol

SESSION 1
-Student enters room-
Researcher: Hi, welkom. Leuk dat je met Pixel wil gaan rekenen. Voor we gaan beginnen zijn er een paar dingen die je moet weten. Alles wordt gefilmd, zodat ik straks kan terugkijken hoe goed jij en de robot bezig zijn geweest, maar deze video wordt over een paar weken weer verwijderd. Wat je ook nog moet weten is dat we eerst even een korte training gaan doen, zodat je kan oefenen met het praten tegen de robot, en weet hoe hij werkt. Dit gaan we nu eerst even samen doen en daarna mag je zelf met de robot gaan rekenen. Dit mag je 2x doen. 1x nu en 1x volgende week. Nadat je klaar bent met sommetjes maken, stel ik je nog een paar vragen, en dan is het alweer klaar. Heb jij nog vragen voor mij?

-Start training-

Researcher: Dat ging goed! Ik ga je nu een blaadje geven, waarop 20 gevoelens genoemd staan. Ik wil dat jij per gevoel aangeeft, hoe jij je nu voelt.
-PANAS meting-
Nu ga ik de robot aanzetten om te gaan rekenen. Onthoud dat je duidelijk moet praten, dan kan hij je goed horen. In principe doe je dit zelf, samen met de robot. Ik ga daar zitten werken, als er nou echt iets is, dan kun je mij om hulp vragen. Is alles duidelijk? Succes!

-Start session-

Researcher: Nou, wat vond je ervan? Dankjewel voor het meedoen vandaag. Voor je teruggaat naar de klas, heb ik nog wat vraagjes voor je.

-PANAS&Questionnaire-

Volgende week ben je weer aan de beurt. Wil je nu -NAME- even voor me halen?

SESSION 2
-Student enters room-
Researcher: Hey, leuk dat je weer mee doet. Heb je er zin in? Vandaag is de laatste keer dat je sommetjes met Pixel gaat oefenen. Weet je nog hoe alles werkt? Als je klaar bent met Pixel, ga ik je weer een aantal vragen stellen en dan mag je weer terug naar de klas. Is alles duidelijk? Heb je nog vragen? Succes! Als er iets is, kun je me altijd even roepen.

-Start session-

Researcher: Nou, dat was het dan. Vond je het leuk? Ik heb nu een aantal vragen voor jou, waarbij ik wil dat je het getal aanwijst dat het best past bij wat jij vindt.

-PANAS&Questionnaire-

Researcher: Ik wil je heel erg bedanken voor het meedoen. Zeg maar dag tegen Pixel, en dan mag je -NAME- even voor me halen als je wil.

K Variables

Column title	Explanation
Number	Student number used in research
School	Name of the school to which the student belongs
Reference #	Reference number for student in the research, used in the paper-work and
data gathering	
Group	Either test or control group
Grade	The CITO grade used at the school for math
Gender	Gender of the student
Age	Age of the student
PanasPre	Weighted result of the PANAS test before the first session, after the training
PanasPost1	Weighted result of the PANAS test after the first session
PanasPost2	Weighted result of the PANAS test after the second session
TestPre	Number of correct answers in the math test taken before the first sessions
TestPost1	Number of correct answers in the math test taken directly after the second sessions
TestPost2	Number of correct answers in the math test taken two weeks after the sessions have ended
TimesTrained	The number of times a student completed the training
TrainingScore	The score given to the training according to the training protocol
1#foot	The number of times the foot buttons have been touched in session 1
1#head	The number of times the head button has been touched in session 1
1#exercises	The number of exercises completed in session 1
1#errorPixel	The number of times Pixel misunderstood a correct answer in session 1
1#tooLate	The number of times a student answered too late in session 1
1#noFoot	The number of times a student forgot to touch the foot button before giving an answer in session 1
1#askedHelp	The number of times a student asked the researcher for help during session 1
1#standUp	The number of times the student stood up from the chair in session 1

1#actSame	The number of times the student acts and moves the same way as Pixel in session 1
1#madNoHear	The number of times the student gets frustrated because Pixel does not hear or understand an answer in session 1
1#madTooSlow	The number of times the student gets frustrated because Pixel is too slow in interacting in session 1
1Notes	Notes taken in session 1
2#foot	The number of times the foot buttons have been touched in session 2
1.0.0	The number of times a student had a correct answer in the first try, in the first category, in session 1
1.0.1	The number of times a student had a correct answer in the second try, in the first category, in session 1
1.0.2	The number of times a student had no correct answer in two tries, in the first category, in session 1
1.1.0	The number of times a student had a correct answer in the first try, in the second category, in session 1
1.1.1	The number of times a student had a correct answer in the second try, in the second category, in session 1
1.1.2	The number of times a student had no correct answer in two tries, in the second category, in session 1
2.2.0	The number of times a student had a correct answer in the first try, in the third category, in session 2
1#total0	The total amount of correct answers in the first try in session 1
1#total1	The total amount of correct answers in the second try in session 1
1#total2	The total amount of no correct answers after two tries in session 1
2#total0	The total amount of correct answers in the first try in session 2
1#rijgen	The amount of mistakes made by forgetting the last step in the lacing method in session 1
1#plus/min	The amount of mistakes by using the incorrect operator in session 1
1#splitsen	The amount of mistakes made by splitting and adding incorrectly in session 1
1#switch	The amount of times a mistake has been made by switching numbers in session 1
2#rijgen	The amount of mistakes made by forgetting the last step in the lacing method in session 2
1Q1	The answer to the first question of the questionnaire in session 1
1Q2	The answer to the second question of the questionnaire in session 1
1Q4.1	The answer to the first subquestion of question 4 of the questionnaire in session 1
2Q4	The answer to the fourth question of the questionnaire in session 2
Tips	The answer to the question whether the students had improvements or tips for Pixel after session 2

L

Exercises

Addition: Passing tens:

4 + 7 = 11	8 + 7 = 15	8 + 5 = 13	7 + 3 = 10
10 + 2 = 12	8 + 4 = 12	8 + 3 = 11	8 + 2 = 10
9 + 8 = 17	1 + 10 = 11	8 + 6 = 14	4 + 10 = 14
5 + 6 = 11	4 + 8 = 12	8 + 10 = 18	

Addition: Tens and units:

44 + 3 = 47	61 + 7 = 68	75 + 1 = 76	46 + 3 = 49
60 + 2 = 62	50 + 9 = 59	61 + 2 = 63	50 + 9 = 59
11 + 6 = 17	11 + 1 = 12	61 + 1 = 62	93 + 2 = 95
74 + 1 = 75	14 + 5 = 19	36 + 3 = 39	71 + 2 = 73
22 + 2 = 22	80 + 3 = 83	64 + 1 = 65	21 + 6 = 27

Addition: Tens and units passing tens:

18 + 55 = 73	15 + 77 = 92	18 + 33 = 51	48 + 33 = 81
57 + 35 = 92	25 + 66 = 91	59 + 39 = 98	38 + 45 = 83
67 + 26 = 93	52 + 29 = 81	18 + 68 = 86	26 + 45 = 71
78 + 13 = 91	42 + 39 = 81	26 + 38 = 64	35 + 63 = 98
38 + 17 = 55	16 + 28 = 44	48 + 48 = 96	24 + 28 = 52
46 + 47 = 93	18 + 24 = 42	34 + 29 = 63	32 + 19 = 51
25 + 46 = 71			

Subtraction: Through tens:

78 - 9 = 69	83 - 5 = 78	15 - 7 = 8	33 - 6 = 27
24 - 8 = 16	62 - 5 = 57	71 - 8 = 63	65 - 7 = 58
27 - 9 = 18	11 - 3 = 8	21 - 7 = 14	31 - 9 = 22
42 - 5 = 37	48 - 9 = 39	52 - 3 = 49	55 - 8 = 47
61 - 5 = 56	66 - 8 = 58	23 - 8 = 15	81 - 6 = 75

Subtraction: Remove tens and add later:

12 - 1 = 11	18 - 6 = 12	24 - 2 = 22	27 - 3 = 24
35 - 2 = 33	38 - 4 = 34	44 - 1 = 43	48 - 6 = 42
53 - 1 = 52	58 - 4 = 54	63 - 2 = 61	68 - 6 = 62
72 - 1 = 71	77 - 4 = 73	85 - 3 = 82	88 - 7 = 81
91 - 1 = 90	93 - 2 = 91	98 - 6 = 92	15 - 3 = 12

Subtraction: Tens minus tens and units:

14 - 12 = 2	18 - 14 = 4	25 - 14 = 11	28 - 16 = 12
33 - 12 = 11	36 - 15 = 21	38 - 22 = 16	42 - 11 = 31
44 - 22 = 22	46 - 33 = 13	53 - 21 = 32	58 - 24 = 34
63 - 11 = 52	69 - 22 = 47	72 - 41 = 31	77 - 34 = 43
84 - 22 = 62	88 - 43 = 45	93 - 11 = 82	95 - 53 = 42

Subtraction: Tens and units through tens:

22 - 16 = 6	46 - 18 = 28	24 - 17 = 7	32 - 23 = 9
35 - 19 = 16	33 - 14 = 19	43 - 25 = 18	47 - 39 = 8
45 - 18 = 27	52 - 33 = 19	55 - 27 = 28	58 - 19 = 39
62 - 35 = 27	64 - 46 = 18	67 - 29 = 38	73 - 36 = 37
77 - 19 = 58	83 - 68 = 15	85 - 47 = 38	94 - 59 = 35

Bibliography

[1] Meer rekenfouten op een tablet (brussel) - de standaard. http://www.standaard.be/cnt/DMF20130628_00639692?utm_source=dlvr.it&utm_medium=twitter&utm_campaign=benjochems. (Accessed on 07/26/2018).

[2] Naoqi apis — aldebaran 2.4.3.28-r2 documentation. http://doc.aldebaran.com/2-4/naoqi/index.html, . (Accessed on 05/28/2018).

[3] Nao robot: characteristics | softbank robotics. https://www.softbankrobotics.com/emea/en/robots/nao/find-out-more-about-nao, . (Accessed on 06/21/2018).

[4] Panas - wikipedia. https://nl.wikipedia.org/wiki/PANAS. (Accessed on 01/29/2019).

[5] T u l e. http://tule.slo.nl/RekenenWiskunde/F-L27.html. (Accessed on 05/23/2018).

[6] Shawn Achor. *The happiness advantage: The seven principles of positive psychology that fuel success and performance at work*. Random House, 2011.

[7] Anton Batliner, Christian Hacker, Stefan Steidl, Elmar Nöth, Shona D'Arcy, Martin J Russell, and Michael Wong. " you stupid tin box"-children interacting with the aibo robot: A cross-linguistic emotional speech corpus. In *Lrec*, 2004.

[8] Sebastian Deterding. Gamification. *Interactions*, 19(4):14, 2012. ISSN 10725520. doi: 10.1145/2212877.2212883. URL http://dl.acm.org/citation.cfm?doid=2212877.2212883.

[9] Zan Gao, Leslie Podlog, and Chaoqun Huang. Associations among children's situational motivation, physical activity participation, and enjoyment in an active dance video game. *Journal of Sport and Health Science*, 2(2):122–128, 2013. ISSN 22132961. doi: 10.1016/j.jshs.2012.07.001. URL http://dx.doi.org/10.1016/j.jshs.2012.07.001.

[10] Rachel Goldman. Using educational robotics to engage inner-city students with technology. *Proceedings of the 6th international ...*, 2003(I):214–221, 2004. URL http://portal.acm.org/citation.cfm?id=1149151{%}5Cnpapers2://publication/uuid/65BF6840-A961-4FCE-B0B8-7C8B88AF6FAF.

[11] Merrill Harmin. *Inspiring active learning: A handbook for teachers.* ERIC, 1994.

[12] John Hattie and Helen Timperley. Review of Educational The Power of Feedback. *Review of Educational Research*, 77(1):81–112, 2007. doi: 10.3102/003465430298487. URL http://rer.sagepub.com/content/77/1/81.

[13] Chien-ming Huang. Generating Effective Social Behaviors for Robots. 2012.

[14] Joris B Janssen, Chrissy C van der Wal, Mark A Neerincx, and Rosemarijn Looije. Motivating children to learn arithmetic with an adaptive robot game. In *International Conference on Social Robotics*, pages 153–162. Springer, 2011.

[15] Mohammad Ehsanul Karim, Séverin Lemaignan, and Francesco Mondada. A review: Can robots reshape K-12 STEM education? *Proceedings of IEEE Workshop on Advanced Robotics and its Social Impacts, ARSO*, 2016-March, 2016. ISSN 21627576. doi: 10.1109/ARSO.2015.7428217.

[16] James Kennedy, Paul Baxter, and Tony Belpaeme. The Robot Who Tried Too Hard: Social Behaviour of a Robot Tutor Can Negatively Affect Child Learning. *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, (801):67–74, 2015. ISSN 21672148. doi: 10.1145/2696454.2696457.

[17] Michael D. Kickmeier-Rust, Eva-C. Hillemann, and Dietrich Albert. Gamification and Smart Feedback. *International Journal of Game-Based Learning*, 4(3):35–46, 2014. ISSN 2155-6849. doi: 10.4018/ijgbl.2014070104. URL http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/ijgbl.2014070104.

[18] Gabriela Kiryakova, Nadezhda Angelova, and Lina Yordanova. Gamification in Education. 2018. ISSN 13300067. doi: 10.4018/978-1-5225-5198-0. URL http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-5198-0.

[19] J M Kraemer. Balans (40) over de strategieën en procedures bij het hoofdrekenen halverwege de basisschool. *PPON-reeks*, 40(40), 2009.

[20] Van Luit, J.E.H. De ontwikkeling van Tellen en getalbegrip bij kleuters. *PO-raad*, pages 1–13, 2013. URL www.rekenpilots.nl.

[21] Christien Janssen Mieke van Groenestijn, Ceciel Borghouts. *Protocol Ernstige RekenWiskunde-problemen en Dyscalculie BAO, SBO, SO*. Van Gorcum, 2011.

[22] Inge Molenaar and Carolien Knoop van Campen. Learning analytics in practice The effects of adaptive educational technology Snappet on students' arithmetic skills. *Proceedings of the 6th International Conference on Learning Analytics & Knowledge - LAK '16*, pages 10–11, 2016. doi: 10.1145/2883851.2883892. URL http://dl.acm.org/citation.cfm?doid=2883851.2883892.

[23] Inge Molenaar, Carolien van Campen, and Karly van Gorp. Onderzoek naar Snappet; gebruik en effectiviteit. *Kennisnet.nl*, page 44, 2016. URL https://www.kennisnet.nl/artikel/leerlingen-presteren-beter-dankzij-slimme-tablet/.

[24] Omar Mubin, Catherine J. Stevens, Suleman Shahid, Abdullah Al Mahmud, and Jian-Jie Dong. a Review of the Applicability of Robots in Education. *Technology for Education and Learning*, 1(1), 2013. ISSN 1916-7008. doi: 10.2316/Journal.209.2013.1.209-0015. URL http://www.actapress.com/PaperInfo.aspx?paperId=43268.

[25] John Norcini. The power of feedback. *Medical Education*, 44(1):16–17, 2010. ISSN 03080110. doi: 10.1111/j.1365-2923.2009.03542.x.

[26] Lucienne van Os et al. Hoe lossen jongens en meisjes aftrekopgaven op? B.S. thesis, 2015.

[27] Seymour Papert. *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc., 1980.

[28] Patrick D Quinn and Angela L Duckworth. Happiness and academic achievement: Evidence for reciprocal causality. In *The Annual Meeting of the American Psychological Society*, volume 24, page 2007, 2007.

[29] Emmy Rintjema, Rianne van den Berghe, Anne Kessels, Jan de Wit, and Paul Vogt. A Robot Teaching Young Children a Second Language: The Effect of Multiple Interactions on Engagement and Performance. (March):219–220, 2018. doi: 10.1145/3173386.3177059. URL http://dl.acm.org/citation.cfm?doid=3173386.3177059.

[30] Softbank Robotics. Discover nao, the little humanoid robot from softbank robotics | softbank robotics. https://www.softbankrobotics.com/emea/en/robots/nao. (Accessed on 05/28/2018).

[31] Martin Saerbeck, Tom Schut, Christoph Bartneck, and Maddy D. Janse. Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, pages 1613–1622, 2010. ISSN 0379864X. doi: 10.1145/1753326.1753567. URL http://dl.acm.org/citation.cfm?id=1753326.1753567.

[32] Namin Shin and Sangah Kim. Learning about, from, and with robots: Students' perspectives. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pages 1040–1045, 2007. doi: 10.1109/ROMAN.2007.4415235.

[33] Ravi Singh, Muhammad Saleem, Prabodha Pradhan, Cristina Heffernan, Neil T. Heffernan, Leena Razzaq, Matthew D. Dailey, Cristine O'Connor, and Courtney Mulcahy. Feedback during web-based homework: The role of hints. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6738 LNAI:328–336, 2011. ISSN 03029743. doi: 10.1007/978-3-642-21869-9_43.

[34] John Swezller and John Sweller. Cognitive Load Theory , Learning Difficulty , and Instructional Design. *Learning and Instruction*, 4:295–312, 1994. ISSN 09594752. doi: 10.1016/0959-4752(94)90003-5.

[35] Richard G Tiberius and Janet Mancini Billson. The social context of teaching and learning. *New directions for teaching and learning*, 1991(45):67–86, 1991.

[36] H Van der Maas, M Straatemeier, and S Klinkenberg. Rekenen in rekentuin: ook voor zwakke rekenaars? *Tijdschrift voor Remedial Teaching*, 17(3):18–21, 2009.

[37] H Van der Maas, S Klinkenberg, M Straatemeier, et al. Rekentuin. nl: Combinatie van oefenen en toetsen. *Examens*, 2010.

[38] Hanneke J.C. van Nuland, Elise Dusseldorp, Rob L. Martens, and Monique Boekaerts. Exploring the motivation jungle: Predicting performance on a novel task by investigating constructs from different motivation perspectives in tandem. *International Journal of Psychology*, 45(4):250–259, 2010. ISSN 00207594. doi: 10.1080/00207591003774493.

[39] P.C.M. Van Soest and W.M. Nolden Van Manen. Protocol ernstige rekenproblemen en dyscalculie groep 5 tot en met 7. *Master Special Educational Needs, Opleidingscentrum Speciale Onderwijszorg Fontys Hogescholen Tilburg*, 2010.

[40] M Verschuren. Feedback in de rekenles. pages 165–177, 2009.

[41] Lev Semenovich Vygotsky. *Mind in society: The development of higher psychological processes*. Harvard university press, 1980.

[42] David Watson, Lee Anna Clark, and Auke Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063, 1988.

[43] Grant Wiggins. 7 Keys to Effective Feedback. *Educational leadership*, 70(1):10–16, 2012. ISSN 00205745.

[44] Daphne Wolsink et al. Hoe lossen basischoollleerlingen aftrekopgaven op? B.S. thesis, 2015.