Noordhoff

UNIVERSITY OF TWENTE.

Brengt je verder

Feedback and drill practice in a digital teaching method

Name:	BSc. R. Davids
Master Program:	Educational Science and Technology
First Supervisor:	Dr. A. M. van Dijk
Second Supervisor:	Dr. H. van der Meij
External Organization:	Noordhoff Uitgevers
External Supervisor:	E. van Vroonhoven
Date Of Submission:	25-08-2020

Table Of Contents

Abstract

E-learning is on the rise. Publishers of learning materials Dutch primary education respond to this by publishing new digital teaching methods. Along with new digital teaching methods, a new way of providing feedback emerges. Noordhoff Uitgevers is such a publisher who published ´Getal & Ruimte Junior´ with four feedback types integrated. Two of these feedback types are the focus of this study: right/wrong feedback and tailored feedback. This study provides information regarding which feedback types works best for enhancing pupils´ achievement and self-efficacy.

A quasi, mixed method, experimental design was utilized to determine which feedback type works best. Pupils attending grade four were appointed to either one of the two conditions: only right/wrong feedback or right/wrong feedback in combination with tailored feedback. The pupils took the Tempo Toets Rekenen, an automated math skills test, as means of a pre-test. After that pupils made several math exercises in Getal & Ruimte Junior, while either seeing only right/wrong feedback or right/wrong feedback in combination with tailored feedback. Next, they filled in a self-efficacy questionnaire. And last, they took a post-test (several math exercises with only right/wrong feedback).

Mixed method analyses showed no significance regarding tailored feedback increasing math achievement in comparison to right/wrong feedback. Neither did it increase pupil’s self-efficacy. Although no evidence was found that tailored feedback increases math achievement or self-efficacy, in talking to the pupils it became clear that they preferred tailored feedback over merely right/wrong feedback.

Digitalization

Midway through the nineties the use of information and computer technology (ICT) increased rapidly in Dutch education (Rubens, 2003). ‘E-learning’ is the given term by 2003 for the use of ICT for learning for both youngsters and people of age. The possibilities of ICT are an opportunity to improve the traditional way of teaching and learning. It was expected that with the aid of the online world learning could be faster, more interactive, and efficient (Dobbs, 2002). Rubens and Admiraal (2003) stated above all that technology looked promising as means to set up learning processes. Communication and interactivity would play a central role. The learner and the learning process would be the core business.

If we take a leap to the present and look inside the classrooms of Dutch primary education, we see that ICT is used for numerous purposes (Kozma, 2003). First, teachers eliminated the old blackboard and now use a Digi board for in class instruction. The Digi board is not only used for showing videos or inserting text, it too is used to show a digitized version of the traditional textbook. Nowadays not only the teachers use ICT in the classroom, the children do this as well. Children attending school in The Netherlands nowadays almost always have access to a computer, laptop, or tablet. These devices are used for learning. Since most publishers digitized most of their teaching methods, pupils are able to solely use a device for learning. No longer actual books are needed for learning or teaching (Shield, 2000).

Along the shift from classical, analogue teaching to digitized teaching the way feedback is delivered changed as well. The new digital teaching method paves the way for more individualized and immediate feedback (Vasilyeva, Pechenizkiy, & De Bra, 2008). Where teachers had to check the work of the pupils, the digital teaching method can do this for them. Pupils are able to receive feedback immediately after finishing an exercise. When pupils are struggling with the exercise, they are able to get feedback. The feedback helps them with that specific exercise.

Moreover, with the Corona virus making its way through the entire population schools are and will remain closed for an uncertain amount of time (World Health Organization, 2020). This means that, now more than ever, teachers and pupils are more or less forced to work online from home. Effective education will now rely on the adaptiveness of the teachers working from home, but more important the digitized teaching methods in which the pupils can on their own and without help from the teacher.

A New Digital Teaching Method

One of the publishers publishing digital teaching methods is Noordhoff Uitgevers (Noordhoff Uitgevers, 2017). In 2017 they published four digitized teaching methods. Teachers instruct on a Digi board and pupils read from a digital textbook. They make exercises, even take tests digitally, while the teacher oversees this from his or her laptop. Learning and teaching could be done entirely digital. In this thesis, the newest digitized teaching method, Getal & Ruimte Junior (GRJR), will be focused upon. In this new digital teaching method, Noordhoff Uitgevers incorporated four ways of presenting feedback, while in previous digital teaching methods less types of feedback were integrated (Noordhoff Uitgevers, 2017).

The choice for the four feedback types in GRJR stems from best practices of the teachers involved in making the new digital teaching method (E. van Vroonhoven, personal communication, May 26, 2017). Several teachers were asked for their input concerning which kind of feedback should be incorporated within GRJR. In consultation with the project leader of GRJR they came to the agreement to incorporate four types of feedback in this new digital teaching method.

Since this agreement was based on the best practices of the teachers involved and not on scientific literature, Noordhoff Uitgevers wanted it researched whether one type of feedback was more effective than the other (E. van Vroonhoven, personal communication, May 26, 2017). Even though feedback has long been regarded as an important factor in learning processes both offline and online, there is little empirical evidence for the effectiveness of feedback regarding mathematics in an online learning environment (Azevedo & Bernard, 1995; Klei, Feskens & Eggen, 2015; Razzaq, Ostrow & Heffernan, 2020).

Thus, the aim of this study, as in accordance with the wishes of Noordhoff Uitgevers, is to clarify and understand which kind of feedback in a mathematical digital teaching method works best for pupils attending grade four in Dutch primary education.

Getal & Ruimte Junior

GRJR is a teaching method for mathematics for primary education. Where previous teaching methods for mathematics were generally provided offline, GRJR is in its entirety digitized (Noordhoff Uitgevers, 2017). Both pupils and teachers are able to work with GRJR solely on a computer, laptop, or tablet. Pupils are still able to receive feedback, although now in a digital form.

Feedback in GRJR is presented in four different ways (Noordhoff Uitgevers, 2017). First and foremost, feedback is provided by means of showing whether the answer is correct or incorrect. The correct or incorrect answers are highlighted in the colours green or red. But also, a red cross appeared for incorrect answers. Second, feedback is presented through a 'hint'. This hint pops up from time to time after an incorrect answer is given. Third, there is elaborated feedback present. Elaborated feedback in GRJR shows the pupils in words how the problem should be dealt with. Fourth and last, another type of elaborated feedback is provided, called tailored feedback. Tailored feedback shows step by step, also in words, how the problem should be solved. However, this time the words are guided by images and audio. These four ways of presenting feedback differ in timing, nature and which process they address.

In this study right/wrong feedback and tailored feedback will be researched. These two types of feedback are chosen mainly, because right/wrong feedback is always present in the digital teaching method. But also, because tailored feedback is the newest feedback type developed by Noordhoff Uitgevers and at the same time the most detailed feedback type present.

Feedback: A Definition

Formative feedback is and has long been regarded an important, crucial practice for learning (Goldin, Narciss, Foltz & Bauer, 2017; Narciss & Huth, 2008; Shute, 2008; Hattie & Timperley, 2007). It is, among others, a strong influencer of one's learning and self-efficacy (Shute, 2008; Hattie & Timperley, 2007). Formative feedback provides the learner with information about one's current state of performance or understanding in order for them to reflect upon it and change their behaviour or thinking accordingly (Goldin, Narciss, Foltz & Bauer, 2017; Narciss & Huth, 2008; Shute, 2008; Hattie & Timperley, 2007). Shute (2008) defined formative feedback as information, which is provided to the learner with the purpose of improving learning by changing one's behaviour or thinking, focusing on the learner as the primary receiver.

Hattie and Timperley (2007) add a sender in their definition and define formative feedback as "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (p. 81). It is stated that different sources of feedback, that is senders, are able to provide feedback with alternative effects on the recipient. A teacher can provide feedback that corrects wrongfully remembered information, a fellow learner can provide feedback that sheds new light on a subject, a parent can provide reassurance in a difficult situation, a website can provide information to help understand a problem, and the recipient itself can look up an answer to check their given response to a problem.

Context is added by Narciss (2017), who states that formative feedback is all post-response information in instructional situations. These instructional situations were previously regarded as offline; while in the present formative feedback can be provided in online learning environments (Goldin, Narciss, Foltz & Bauer, 2017; Leibold & Schwarz, 2015). Now formative feedback is not only provided in written text, but, through information technology, also by audio, video, online programs/games or live in an online meeting or MOOC (Goldin, Narciss, Foltz & Bauer, 2017; Leibold & Schwarz, 2015).

In this study, formative feedback is provided in an online setting; the digitized teaching method Getal & Ruimte Junior (Noordhoff Uitgevers, 2017). The feedback is pre-programmed by the developers, however adapted to each exercise. It aims to correct and instruct learners how to perform better on math exercises. Thus the definition fitting to this situation, regarding right/wrong and tailored feedback, and used in this study is: formative feedback is information provided by an online agent in order to modify one's understanding regarding math, in this case long multiplication.

Feedback In Getal & Ruimte Junior

Formative feedback could take on various functions and miscellaneous formats or types. Narciss (2008) provides a list of types of feedback, which are widely used among researchers and practitioners. One type of feedback on that list is knowledge of results, which indicates whether the learner's input was correct or incorrect without providing any additional information nor the correct answer (Narciss, 2008). Error flagging is such knowledge of results feedback, which too does not provide further information nor the correct answer. However, it does show the learner the location of the incorrect response (Klei, Feskens & Eggen, 2015; Shute, 2008). This type of feedback intents to stimulate the correct recollection of facts. It, however, could be seen as a trial-and-error procedure. Looking at the feedback provided by GRJR, one could see that right/wrong feedback is similar to knowledge of results. The right/wrong feedback shows where the learner made an error but does not show the correct answer nor does it provide additional information. This could be seen as error flagging or multiple-try feedback, since learners get two tries for task completion while receiving solely knowledge of results and no knowledge of correct response before receiving tailored feedback (Klei, Feskens & Eggen, 2015; Narciss, 2008; Shute, 2008).

Also mentioned is knowledge of correct response, which provides the learner with the correct answer. Yet another type that is brought up is elaborated feedback. Elaborated feedback is feedback that states the correct response while at the same time providing additional information. Elaborated feedback can take many forms and thus can entail many different meanings (Kleij, Feskens & Eggen, 2015; Shute, 2008). Elaborated feedback could take the form of hints, worked examples, additional information or extra study materials. Elaborated feedback, in general, focuses on the correct response, can explain why an incorrect answer is wrong and can show what the correct answer should be. Knowledge of results or knowledge of correct response, however, often co-occur with elaborated feedback. With modern technology formative feedback could be implemented in a wider range of possibilities. Such an implementation is informative tutoring feedback, which does not provide the correct answer at first, but provides the learner with important information towards task completion (Narciss & Huth, 2006). The feedback in GRJR, known as tailored feedback, is a combination of elaborated feedback and knowledge of correct response. Tailored feedback provides the learner with information how to complete the exercise, while showing the answer in the end. However, tailored feedback does not show the precise point where the learner has made the mistake.

In general, the aim of formative feedback is, according to Hattie & Timperley (2007), to reduce the gap between current and future understanding, performance, and goals. Shute (2008) skips the gap and states that the preeminent goal of feedback is to broaden one’s skills, knowledge and understanding. As do Leibold & Schwarz (2015), who state that formative feedback identifies strengths and provides the learner with information how one could improve and develop. Shared among the authors is the purpose of improving one’s understanding and performance. The purpose of the feedback focused upon in GRJR, is to provide the learner with information about his or her performance and understanding. Right/wrong feedback helps the learner signal incorrect inputs, while tailored feedback promotes the use of the correct strategies for solving math problems.

Feedback could take on two functions: directive and facilitative (Shute, 2008). Directive feedback points out to the learner what needs revision or fixing. Knowledge of results and knowledge of correct response are formative feedback types that take on the function of directive feedback, i.e. they show the learner where the mistake is. A third function is mentioned by Klei, Feskens & Eggen (2015) and Leibold & Schwarz (2015). This third function is the corrective function, which provides the learner with information in order to correct the error that was made. In GRJR right/wrong feedback takes on the directive and corrective function. It points out to the learner where mistakes were made by showing the colour red, and thus directs the learner to the exercise in which the mistake was made. In terms of the corrective function, right/wrong feedback does provide information in order to correct the error. However, the extent to which the colour red could be useful information for error correction is limited. Facilitative feedback is less specific in that it provides the learner with remarks and notes, which guides the learner in the process of revision (Shute, 2008). A feedback type that takes on such a facilitative feedback function is elaborated feedback, which could provide hints and tips for revision. It provides the learner with information on how to solve the problem, without showing the solution. However, tailored feedback in GRJR does not provide guidance for revision as it merely states the rules on how to solve the math problem, and in turn expects the learner to extract the specific useful information for the error that was made.

Four major levels of formative feedback are suggested (Hattie & Timperley, 2007; Klei, Feskens & Eggen, 2015), and the effectiveness of formative feedback differs to which level the formative feedback is directed. The task or product is the first level at which formative feedback could be focused upon, which shows the learner whether the attempts for task completion are correct or incorrect. Formative feedback at the level of task or product may also contain directions for the learner. It prompts the learner to collect extra or different information for task completion. The second level of formative feedback is formative feedback aimed at the processes used during and for task completion. For example, information processing or learning processes needed to understand the task. In GRJR right/wrong feedback is clearly focused on the level of the task. The learner receives feedback showing correct and incorrect responses. It, however, does not provide the learner with directions, as it does not prompt the learner to collect further information for task completion. Whether right/wrong feedback influences the learner at the process level is arguable. It does not provide comments on how to alter processes, it however could inspire the learner to change thinking processes. Tailored feedback, on the other hand, does focus on the level of processes. It prompts the learner to rethink their learning processes regarding to rules for task completion. Ideally, learners compare the rules for task completion they used to the rules for task completion they should use according to the digital teaching method.

Formative feedback may either be presented immediately or delayed (Shute, 2008). Immediate formative feedback is presented to the learner right after completing a task or responding to an item**.** Delayed formative feedback is relative to immediate formative feedback. It therefore is defined as feedback provided to the learner minutes, hours or even weeks after task completion. Right/wrong feedback in GRJR could be considered immediate formative feedback. After entering their answers pupils click ‘check’ and receive immediate feedback on correct or incorrect responses. Tailored feedback, however, could be considered delayed feedback. Tailored feedback is only shown after the pupil takes two tries for task completion and still fails to complete the task successfully. In this case delayed means after two tries for task completion.

Feedback Effectiveness For Learning

Bloom (1984) as well as Kleij, Feskens & Eggen (2015) report that one-to-one tutoring is a form of instruction and feedback that is the most effective compared to other forms of tutoring. One-to-one tutoring is immediate, adaptive, and differentiating. For example, a teacher could provide immediate feedback to a pupil when he or she gets confronted with a misunderstanding. This feedback could be adapted and differentiated to the needs of the learner. In an online learning environment such one-to-one tutoring is almost impossible, so different feedback in different forms are applied. These differing online forms of feedback are effective in different settings for different processes.

Klei, Feskens & Eggen (2015) introduce three studies providing an overview of the effectiveness of feedback in an online learning environment on learning outcomes: Azevedo and Bernard (1995), Jaehnig and Miller (2007), and Van der Kleij et al. (2011). From these studies the following is concluded: knowledge of results seems to be ineffective for learning. Knowledge of correct response looks moderate effective for lower order learning processes. Elaborated feedback seems to yield to highest effective results regarding both lower and higher learning processes. In the study of Jaehnig and Miller (2007) the effectiveness of neither immediate nor delayed feedback over one or another has been proven, although no clear definitions for immediate or delayed feedback were provided.

Narciss & Huth (2006) report that some beneficial effects of bug-related tutoring feedback on achievement are found in some studies. These studies selected components form bug-related tutoring feedback reasoned from cognitive task and error analyses and applied these components to a multiple try feedback algorithm. They also state that task-specific tutoring feedback yield positive effects on performance and motivation.

“Feedback has no effect in a vacuum; to be powerful in its effect, there must be a learning context to which feedback is addressed” (Hattie & Timperley, 2008, p. 82). Formative feedback at the level of task yields the most effects when it addresses wrongful understanding, rather than a gap in knowledge (Narciss & Huth, 2006). However, it does usually not transfer to other tasks. Too much and too specific formative feedback at the level of task may result in trial-and-error procedures. Formative feedback aimed at the processes seems to be more effective for increasing deeper learning processes in comparison to formative feedback aimed at the level of task. Formative feedback aimed at the processes is most effective when it stimulates learners to detect errors and prompts them to look for further information. Formative feedback directed at regulatory or metacognitive processes is effective in the sense that it promotes further engagement with or investment in the task. It increases self-efficacy and motivation, and it promotes a sense of deservedness. Formative feedback directed towards the self is not effective per se, it is rather often present in feedback and used instead of formative feedback at the level of task, formative feedback aimed at the processes, or formative feedback directed at regulatory or metacognitive processes. It must be noted that the effectiveness regarding the four levels of feedback from Narciss & Huth (2006) does not entail feedback in an online learning environment, but rather feedback provided by the teacher to the pupils in the classroom.

Right/wrong feedback in GRJR, according to literature, seems to be ineffective for learning as knowledge of results, but effective as formative feedback at the level of task. Thus, literature shines a contradictory light on right/wrong feedback. However, literature has a more consentient view on tailored feedback. Tailored feedback should be more effective for both lower and higher order learning, as does it promote the motivation of the learners.

Feedback Effectiveness For Self-efficacy

Feedback is not only considered as an important practice for learning, it can also have major influences on self-efficacy (Shute, 2008; Hattie & Timperley, 2007). Self-efficacy is defined by Wood & Bandura (1989), as “beliefs in one’s capabilities to mobilize the motivation, cognitive resources, and courses of action needed to meet given situational demands” (p. 408). One’s self-efficacy is derived from four possible sources (Bandura, 1997). One of those sources is successfully completing a task, a success experience, which may increase self-efficacy. An unsuccessful completion of a task, on the other hand, may cause self-efficacy to decline. Self-efficacy is perhaps more influenced by the allocation of success or failure than the success or failure itself (Hattie & Timperley, 2007). If one is unable to relate feedback to their failure, in other words the feedback is unclear as of why one is doing something wrong, it can have serious consequences for their self-efficacy.

Feedback with little task-related information, like right/wrong feedback, rarely enhances self-efficacy or understanding of the task (Hattie & Timperley, 2007; Wang & Wu, 2008). While feedback providing information about whether one is using the right strategy to complete a task, like elaborated feedback, improves self-efficacy (Wang & Wu, 2008; Hattie & Gan, 2011). Elaborated feedback is helpful in understanding how to complete a task and in turn developing the confidence of having the capability to complete the task successfully, thus increasing self-efficacy (Wang & Wu, 2008). Since right/wrong feedback does not specify why one made a mistake it fails in developing an understanding of the right strategies to use and in turn fails to increase self-efficacy (Hattie & Timperley, 2007).

Feedback Effectiveness For Children

Immediate feedback is found to be more effective for children’s problem solving in mathematics (Fyfe & Rittle-Johnson, 2016). Children who received immediate feedback were less likely to use wrong strategies on the next problems and instead tried to use alternate strategies. In their study the feedback given to the children consisted of the right strategy to use for problem solving without giving the correct answer (Fyfe & Rittle-Johnson, 2016). In other words, the children received elaborated feedback without knowledge of results. These results are endorsed by an earlier study by Brosvic, Dihoff, Epstein & Cook

(2006). They found that immediate, affirmative, and corrective feedback yielded the most effects regarding the acquisition and retention of strategies used while solving math problems.

It has to be noted that prior knowledge moderates the impact of given feedback, whether that feedback is focused on used strategies or correctness of the answer did not matter (Fyfe & Rittle-Johnson, 2016; Fyfe, Rittle-Johnson, & DeCaro, 2012; Rittle-Johnson, Fyfe, Hofer, & Farran, 2016). Children with low prior knowledge benefited more from feedback, resulting in higher test-scores on post-tests. On the other hand, feedback was not beneficial for children with high prior knowledge, resulting in lower test-scores on post-tests (Fyfe & Rittle-Johnson, 2016; Rittle-Johnson, Fyfe, Hofer, & Farran, 2016).

Research Questions

The research question follows logically reading the literature and the wishes of Noordhoff Uitgevers. The main research question guiding this research is stated as follows: how does elaborated feedback compared to right/wrong feedback in a digital teaching method influence fourth grade pupils regarding math achievement and self-efficacy?

In order to answer the main research question three sub questions are stated. Since it will be the pupils first time seeing the new elaborated feedback and using the new digital teaching method it is of importance to see whether there is an effect visible during the main part of the research. This leads to the first sub question:

1. How does feedback in a digital teaching method affect the pupils' math achievement while practicing the math exercises?

The remaining two sub questions follow naturally:

2. How does feedback in a digital teaching method affect the pupils' self-efficacy?
3. How does feedback in a digital teaching method affect the pupils' math achievement?

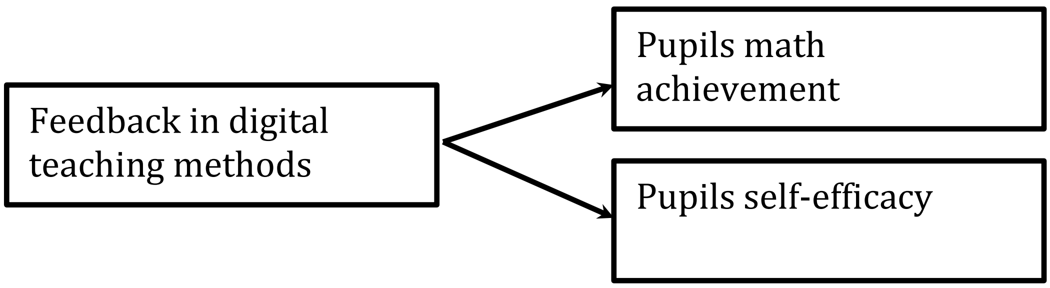


Figure 1: conceptual model for this study with the independent variable 'feedback in digital teaching methods', which positively influences the dependent variables 'pupil achievements in math' and ' pupils' self-efficacy'

Method

Design

This study had a mixed method design and was conducted according to a quasi-experimental design, i.e. two groups made math exercises in GRJR where the control group (N=32; 14 female, 18 male) received right/wrong feedback and the experimental group (N=33; 17 female, 16 male) received both right/wrong feedback and tailored feedback. Both groups took a demographic questionnaire, the Tempo Toets Rekenen, a self-efficacy questionnaire, a post-test and made math exercises in GRJR. Schools were chosen according to availability and willingness. Randomisation of participants was only possible within their respective schools; the pupils were randomly divided between the control and experimental group within their own school. Since schools voluntarily participated and randomisation between schools was not possible, a true experimental design was not feasible.

Participants

In this study, a total of N=60 participants were focused on. The participants were fourth grade pupils, aged 9-11, of which 31 were female (M_{age} = 9.65, SD_{age} = 0.55) and 29 were male (M_{age} = 9.88, SD_{age} = 0.54). The pupils originated from four schools: 23 from school one, six from school two, 24 from school three, and seven from school four. Besides attending grade four in Dutch primary education, being educated in math, and having access to a computer, laptop or tablet, there were no other inclusion or exclusion criteria. However, the schools needed to fulfil one requirement: they needed to be registered in the customer base of Noordhoff Uitgevers.

Instruments

Demographic Questionnaire

At the start of the experiment the pupils were asked to fill in a demographic questionnaire on paper, see appendix C. The demographic questionnaire contained two questions. The pupils were asked to state their sex and date of birth. Instead of asking names the pupils were given a reference number, which is used to identify the work of the pupils and to assure anonymity.

Tempo Toets Rekenen

Pupils took the Tempo Toets Rekenen on paper (a standardized test for basic automated math skills, see appendix D) to determine a baseline for basic math skills (De Vos, 1992). The baseline acted both as a factor to determine the equality between pupils and as reference point for the outcomes of the actual experiment.

The test for basic automated math skills consisted of five columns of 40 math exercises. Four of these columns were exercises for addition, retraction, duplication, and division of natural numbers. The fifth column was a combination of these four basic fundamental operations. The pupils had one minute per column to answer as many questions, as quickly and correctly as possible. All correct answers were counted, and the total amount of correct answers made up their final score.

Getal & Ruimte Junior

In GRJR pupils made math exercises during which they received feedback. In this study we focused on exercises for long multiplication of natural numbers. After entering an answer, pupils clicked on ‘check’ after which the program checked the answer and provided the pupils with right/wrong feedback. See appendix A for a visual explanation of the right/wrong feedback. The right/wrong feedback consisted of green highlighted numbers when the answer was correct. Red highlighted numbers and a red coloured cross when the answer was incorrect. However, when the answer was correct the program only briefly showed the green highlighted answer and automatically generated a new exercise.

After the pupils checked their answer and it was incorrect, pupils clicked on ‘try again’ to try and correct their mistake. After trying to correct their mistake pupils clicked on ‘check’. The program checked the answer and again provided the pupils with right/wrong feedback. This time, when the answer was correct, pupils advanced to the next exercise. If, however, the answer was again incorrect, a new screen appeared with tailored feedback. See appendix A for a visual explanation of tailored feedback in GRJR. In this new screen tailored feedback is presented to the pupils. The specific exercise, which the pupil failed to complete correctly twice, was explained step by step in words and images, guided by audio.

During the experiment pupils made up to 15 math exercises in the digital teaching method. However, a pupil who completed every math exercise correctly would only make a maximum of nine math exercises. The digital teaching method operated in such a manner that after every three math exercises it automatically changed the difficulty level of the subsequent exercises. Three different difficulty levels were available in GRJR: low, moderate, and high. The digital teaching method switched automatically between these depending on the performance of the pupils. Throughout the experiment, the parameters for the different difficulty levels remained the same. After completing up to 15 math exercises the pupils would normally advance to the next assignment in GRJR, however this was not a part of the experiment.

The difference between the experimental and control group was the presented feedback. Pupils in both groups received right/wrong feedback after completing math exercises, either correct or incorrect. The experimental group received tailored feedback in addition to the right/wrong feedback. After failing twice to complete the math exercise correctly, a pop-up screen appeared with the tailored feedback.

The math exercises for both groups were automatically generated according to the following parameter: [110 to 450] x [2 to 9] = [100 to 4000]. The parameter was the same for every pupil and every math exercise. This parameter was chosen in consultation with the publisher and project manager of GRJR, and after analysing the results of the pilot, see appendix E.

The choice for these particular math exercises was very practical. Noordhoff Uitgevers did not have any other exercises with tailored feedback available. Only for these exercises the tailored feedback was on par with the set criteria. So, to test whether tailored feedback had any effect these exercises with completed tailored feedback were chosen.

Self-efficacy Questionnaire

After the experiment, the pupils filled in a self-efficacy questionnaire on paper. The questionnaire contained ten statements from the General Self-efficacy Scale (Schwarzer & Jerusalem, 1995) to measure the self-efficacy of the pupils. An additional five statements (statements 3, 7, 10, 12 and 15) were added to measure pupils' self-efficacy regarding received feedback in GRJR, see appendix F.

This meant that two types of self-efficacy were being measured. On one hand general self-efficacy, for example 'I always succeed in solving difficult problems, if I put in enough effort.' (translated from Dutch). And on the other hand, self-efficacy regarding math achievement, for example 'I am certain that I am able to complete the exercises without any help.' (translated from Dutch). Analysis showed an acceptable Cronbach's alpha of 0.76. The general self-efficacy questionnaire showed an acceptable Cronbach's alpha of 0.78, while the questionnaire for self-efficacy regarding math achievement showed a questionable Cronbach's alpha of 0.63.

Post-test

Next, pupils completed ten math exercises on paper, see appendix H. The post-test was the same for every student. Pupils took the post-test to establish whether the presented feedback had an effect, i.e. whether the pupils performed better on these compared to the previous math exercises. While completing these math exercises pupils received no feedback at all. The math exercises in the post-test were automatically generated according to the same parameter used for the math exercises used in the digital teaching method: [110 to 450] x [2 to 9] = [100 to 4000]. The math exercises in the post-test and in Getal & Ruimte Junior are essentially the same, but with different natural numbers. The maximum score on the post-test was ten.

Procedure

Schools in the customer base of Noordhoff Uitgevers were contacted by phone and/or e-mail (see appendix B for the e-mail) and asked if they wanted to participate in this study. Four schools signed up for the study, after which they were contacted by phone to provide them with more information. The schools shared the contact information of the fourth-grade teachers to set up the experiment. An informed letter of consent (see appendix G) was sent to all parents of the pupils in order to provide them with information about the study and the anonymity of the pupils and the results.

The lead researcher was present during the experiment and supervised it together with the teacher of the pupils. First, the pupils needed to read an introduction and the instructions. After this they took the demographic questionnaire. Next, the pupils made the Tempo Toets Rekenen. After that, the pupils made the math exercises in the digital environment from GRJR on their device. They did this alone and without any help from others. The control group made the exercises without tailored feedback, the experimental group with tailored feedback. Following the math exercises the pupils filled in the self-efficacy questionnaire. Last, the pupils took the post-test. Finishing the post-test, they either handed in the forms or pressed send on their device and continued with their daily tasks set by the teacher.

Depending on the school pupils either had their own device, i.e.: either a tablet or a laptop, or pupils took the tests on a computer provided by school. During the experiment both the lead researcher and the teacher present made sure pupils would not collaborate, cheat, or talk. After finishing the experiment pupils resumed their daily tasks. The experiment took no more than 45 minutes to an hour. The results of the experiment were provided to the schools and their pupils in the form of a small report.

Context

While the setting and procedure of the experiment were the same, the context was very specific and differed from school to school. GRJR was a brand-new digital teaching method. Not one pupil or teacher had seen the digital environment, nor were they familiar with the way it presented long multiplication exercises. Since no school used GRJR, this meant that pupils from different schools were taught different tactics on how to solve long multiplication problems. In addition to different tactics, the tempo of which a class progresses through a teaching method differs per school. This meant that some pupils already had these kind of math exercises explained to them, while other pupils had not. Pupils from different schools, thus, did not possess the same prior knowledge on how to tackle these kind of math exercises.

Analysis

Once the data was collected the results of the control group and experimental group were compared on how feedback affected pupils' math performance and self-efficacy. First, an independent samples T-test was performed to establish differences on the demographics between the two groups, Chi-square tests was used to determine a difference regarding gender, age and completing math exercises in GRJR. Next, several independent samples T-tests were performed in order to establish whether right/wrong feedback or elaborated feedback were of influence. Last, an ANCOVA was used the determine if gender was a predictor for the outcome of the post-test.

Results

Equality Between The Experimental And Control Group

Gender

The distribution of the female and male participants across the experimental and control group are shown in table 1. A Chi-square test shows no significant outcome, i.e. no association was found between the experimental and control group $X2(1)=0.60$, $p=.44$. Neither an association was found between gender and the post-test $X2(7)=9.47$, $p=.22$.

	Sex					
Condition	*Female*	*%*	*Male*	*%*	*Total*	*%*
Control	14	46.7	16	53.3	30	50.0
Experimental	17	56.7	13	43.3	30	50.0

Table 1: distribution of gender

Age

Table 2 shows the distribution between the experimental and control group regarding age. The independent samples T-test shows that a significant difference between the experimental and control group is present t(58)=2.77, p=.01.

	Age					
Condition	*9*	*%*	*10*	*%*	*11*	*%*
Control	5	16.7	22	73.3	3	10.0
Experimental	13	43.3	17	56.7	0	00.0

Table 2: distribution of age

School

The Chi-square test shows no significant results for the association between the school the pupils are attending and whether they are in the experimental or control group X2(3)=0.19, p=.98. The same applies to the association between the variable school and the results on the post-test X2(21)=25.95, p=.21. The distribution of the participants across the four different schools is shown in table 3.

	School							
Condition	*School 1*	*%*	*School 2*	*%*	*School 3*	*%*	*School 4*	*%*
Control	12	40.0	3	10.0	12	40.0	3	10.0
Experimental	11	36.7	3	10.0	12	40.0	4	13.3

Table 3: distribution of schools

Tempo Toets Rekenen

Before the experiment, to establish a baseline in math performance, the pupils took a Tempo Toets Rekenen. The independent samples T-test shows no significant difference between the experimental and control group t(58)=1.81, p=.08. The mean and standard deviations are shown in table 4.

	Tempo Toets Rekenen	
Condition	*M*	*SD*
Control	116.43	29.33
Experimental	103.90	24.10

Table 4: means and standard deviations of the Tempo Toets Rekenen. The maximum score is 200

The Influence Of Feedback On The Exercises Within Getal & Ruimte Junior

To answer the main research question three sub-questions are stated. The first sub-question aims to see whether feedback has an influence on math achievement during the completion of the math exercises within GRJR. To answer that question a Chi-square test was performed on the total label counts of the fifteen exercises, which were available within GRJR for the pupils to make. No significant association was found between feedback and how pupils performed on these exercises $X^2(3)=2.94$, p=.40. In table 5 the total label counts of the exercises within GRJR are shown.

	Math exercises label counts				
Condition	*0*	*1*	*2*	*3*	*Total*
Control	93	38	62	257	450
Experimental	101	28	57	266	450

Table 5: label counts of the exercises within GRJR. The different labels have the following meaning: 0= item was not made, 1= item incorrect, 2= item correct in two, and 3=item correct in one

The Influence Of Feedback On Self-efficacy

The second sub-question aims to answer whether feedback has any influence on the pupils' self-reported self-efficacy. The self-efficacy questionnaire can be split in two parts: general self-efficacy and feedback self-efficacy. The independent samples T-test shows no significant results for the self-efficacy questionnaire overall, nor for the general part t(58)=0,03 p=.98 or the feedback part t(58)=0.15, p=.88.

	Feedback self-efficacy		General self-efficacy	
Condition	*M*	*SD*	*M*	*SD*
Control	3.45	0.59	3.06	0.42
Experimental	3.43	0.43	3.06	0.51

Table 6: means and standard deviations of the self-efficacy questionnaire. The different scores have the following meaning: 1= totally disagree, 2=slightly disagree, and 3=slightly agree, and 4=totally agree

The Influence Of Feedback On Math Achievement

The last sub-question tries to answer whether feedback has any influence on math achievement on the post-test. The independent samples T-test reports no significant results regarding the effect of feedback on math achievement t(58)=-0.84, p=.41. A one-way between subjects ANCOVA was calculated to determine the effect of feedback on the post-test controlling for age. The results of the test shows that age predicts the outcome on the post-test significantly F(1,56)=5.84, p=.02. However, even when controlling for age there is no significant difference regarding the scores on the post-test between the experimental and control group F(1,56)=0.58, p=.45.

	Post-test	
Condition	*M*	*SD*
Control	8.33	2.34
Experimental	8.77	1.59

Table 7: means and distribution of the scores on the post-test. The maximum score is ten

Discussion And Conclusion

The aim of this study is to find out whether right/wrong feedback or tailored feedback improves math achievement and enhances self-efficacy, and if one type of feedback is more effective than the other. Although right/wrong feedback and tailored feedback take on different functions (directive, corrective, and facilitative), are directed to different levels within the learning process (level of the self in contrast to level of the task), and differ regarding at what time they are presented (immediate upon checking the answer opposed to delayed; after two incorrect answers), no tests or questionnaires covering these variables were implemented in this research (Klei, Feskens & Eggen, 2015; Narciss 2008; Narciss & Huth, 2006; Shute, 2008). This research thus focuses on the effectiveness of two types of feedback: namely right/wrong feedback versus tailored feedback. Other aspects of feedback mentioned were not researched, this implies that no well-grounded conclusions can be made about these aspects.

To answer the presented research question, mixed method analyses were conducted, including different types of quantitative research. Using a quasi-experimental design, the control group (right/wrong feedback) was compared to the experimental group (tailored feedback) on a pre- and post-test of the digital teaching method GRJR. Next to math achievement, it was researched whether feedback has influence on self-efficacy of the pupils. This was done by presenting the pupils with a self-efficacy questionnaire.

Pupils in both groups made a comparable amount of math exercises, which means they saw a comparable amount of feedback in GRJR. Looking at the literature, it was expected that the experimental group would perform better after being provided with tailored feedback (Fyfe & Rittle-Johnson, 2016; Fyfe, Rittle-Johnson, & DeCaro, 2012; Rittle-Johnson, Fyfe, Hofer, & Farran, 2016). Looking at the data, it can be concluded that after being presented with tailored feedback, it is not a given that a pupil completes the following exercise correctly. Regarding the influence on math achievement of either type of feedback, it can be concluded that in GRJR both types of feedback did not raise math achievement; both groups scored comparable on the exercises within GRJR and the post-test.

As for the influence of right/wrong feedback or tailored feedback on self-efficacy, no significant results were found. It cannot be concluded that either type of feedback raises the self-efficacy of the pupils. Looking at the data one can see a slightly higher reported self-efficacy regarding the math exercises in the experimental group. This may be a hint that tailored feedback does raise self-efficacy, however no significant results during this experiment were found that support this claim. These findings are contradictory to available literature that states that right/wrong feedback rarely increases self-efficacy, while elaborated feedback does increase self-efficacy (Chang & Lam, 2010; Hattie & Gan, 2011; Hattie & Timperley, 2007; Wang & Wu, 2008).

Limitations

Although Noordhoff Uitgevers claims that GRJR is an adaptive digital teaching method, in its core it cannot be seen as fully adaptive. The difficulty of the exercises differentiates between three levels. There is a baseline for moderate difficulty. When the pupil performs poorly, the digital teaching method switches to a lower difficulty. Does the pupil perform well? Then it switches to a higher difficulty. However, the digital teaching method never switches to another difficulty level than the three aforementioned. On top of that, it only switches levels after three exercises and not after every single exercise. So, the GRJR only adapts to the pupils' needs and level to a certain degree.

The Tempo Toets Rekenen was used to determine a baseline for the pupils' basic math skills (De Vos, 1992). It could be argued that the Tempo Toets Rekenen is not the best way to determine a baseline in basic math skills. First because basic math skills consist of three domains: numbers and operations, ratios, fractions and percentages, and measuring and geometry (Leerlijnen, E. D., 2008; Treffers, A., & Moor, E. D., 1990). Meanwhile the Tempo Toets Rekenen only tests for one domain, thus the results show an incomplete representation regarding pupils' basic math skills. Second, in addition to covering one domain the Tempo Toets Rekenen incorporates a time limit. This time limit could put time pressure on the pupils and influence their state of mind and thus their performance on the test. Last, due to insufficient research little to no evidence is available proving the reliability and validity of the Tempo Toets Rekenen. Therefore, the reliability and validity of the test are questionable (Evers et al., 2009-2012).

Another limitation rests in the research itself; only 15 exercises were to be made during which they were able to experience the feedback in GRJR, while the pupils who did not make any mistakes only made nine exercises. This means that on one hand there were limited moments pupils saw the feedback and could really learn from it. On the other hand, it is possible a pupil did not see any feedback at all. Generally, it is suggested that more feedback is better to improve learning (Salmoni, Schmidt, & Walter, 1984). The limited number of exercises combined with the possibility of even less feedback being presented, means that high effects of tailored feedback in GRJR on pupils' math achievement were not to be expected.

Self-efficacy is influenced by the user experience of a well performing programme, or digital teaching method in this case (Kearsley, 2002). For example, the digital teaching method contained spoken text. Not all pupils could experience this however, because some of them did not bring any earphones to school. This meant that only some pupils could experience the digital teaching method to its full extent. Not being able to listen to the provided audio could mean a decline in self-efficacy. Another example is accessibility; Wi-Fi access (Kearsley, 2002). In order for the digital teaching method to work properly, it needs a stable and reasonably fast internet connection. However, not every schools' Wi-Fi-network could handle this many laptops connecting to the network at the same time. This resulted in pupils not being able to log in, crashes and slow loading pages. Pupils became frustrated and agitated, which did not help the pupil's motivation to continue with the experiment. And in turn possibly influenced pupils' math achievement.

Positive reinforcement is known to be a factor to increase motivation and self-efficacy (Stipek, 1993). In GRJR there is no such thing, besides green highlighted numbers when completing the exercise correctly. Pupils are not able to collect coins for example or even points. Pupils are not reinforced to perform better or reach for a goal; no extrinsic motivation is present. The digital teaching method contains many exercises in a row, which one can check immediately, but does not offer anything else.

In order for the pupils to reach the right page within the digital teaching method they had to follow certain steps. Nine separate steps to be exact. Although these steps were in a clear and orderly fashion presented, many pupils failed to follow these steps exactly. This resulted in pupils landing on the wrong page or not being able to start with the exercises. In turn, the researcher and teachers present had to help the pupils reach the correct page. This took some extra time and it disrupted the calm atmosphere, possibly influencing math achievement as well as pupils' self-efficacy.

Implications For Future Research

In future research, some conditions should be kept in mind. Conditions for sound research should be perfect and stay the same throughout the research, however technology is not perfect nor stable. Researchers should strive for optimal technological conditions, such as the best and fastest internet connection possible when using a digital teaching method. Without a stable and fast internet connection, the digital teaching method could not run optimal if at all. This caused pupils' motivation to decline. Besides the internet connection, the device on which the digital teaching method runs does matter as well. One school in this study used significantly older devices in comparison to the other schools. This caused the digital teaching method to run less smooth and steady overall. For future research it is recommended that every pupil uses the same device, or devices with similar specifications.

During this research, an altered version of the self-efficacy questionnaire was used. Although the statements were simplified in collaboration with several primary school teachers, some questions still proved to be too difficult for the pupils to comprehend. In future research the self-efficacy questionnaire should be even more simplified, or pupils' self-efficacy should be measured in a completely different way.

In general, the results showed no significant effect on the influences of elaborated feedback on math achievement or self-efficacy. Still, the overall scored self-efficacy rating of the pupils in the experimental group was slightly higher compared to the control group. This could mean pupils would be more empowered by more elaborated feedback. Interacting with the pupils during the study revealed that pupils preferred tailored feedback over plain right/wrong feedback. This corresponds to what literature states about elaborated feedback: that it is most effective for learning and enhancing self-efficacy. What can be taken away from this is that publishers should invest in making digital teaching methods not only more adaptive, but also more engaging for the learner. Digital teaching methods should provide more extrinsic motivation and elaborated feedback tailored to the needs of the learner.

Reference List

Alvarez, I., Espasa, A., & Guasch, T. (2012). The value of feedback in improving collaborative writing assignments in an online learning environment. *Studies in Higher Education*, *37*(4), 387-400.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, 13(2), 111-127.

Ashford, S. J., Blatt, R., & Walle, D. V. (2003). Reflections on the looking glass: A review of research on feedback-seeking behavior in organizations. *Journal of Management*, *29*(6), 773-799.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74.

Chan, J. C., & Lam, S. F. (2010). Effects of different evaluative feedback on students' self-efficacy in learning. *Instructional Science*, *38*(1), 37-58.

Ciampa, K. (2014). Learning in a mobile age: an investigation of pupil motivation. *Journal of Computer Assisted Learning*, *30*(1), 82-96.

Clariana, R. B., Ross, S. M., & Morrison, G. R. (1991). The effects of different feedback strategies using computer-administered multiple-choice questions as instruction. *Educational Technology Research and Development*, *39*(2), 5-17.

De Vos, T. (1992). *Tempo-Test-Rekenen: test voor het vaststellen van het rekenvaardigheidsniveau der elementaire bewerkingen (automatisering) voor het basis-en voortgezet onderwijs: handleiding*. Berkhout.

Dobbs, K. (2002). The State of Online Learning – What the Online World Needs Now: Quality. In: Rosset, A. (2002). *The ASTD E-learning Handbook*. New-York: McGraw-Hill, 357-364.

Espasa, A., & Meneses, J. (2010). Analysing feedback processes in an online teaching and learning environment: An exploratory study. *Higher Education: The International Journal of Higher Education and Education Planning*, 59, 277-292. doi: 10.1007/s10734-009-9247-4.

Evers, A., Egberink, I. J. L., Braak, M. S. L., Frima, R. M., Vermeulen, C. S. M., & Vliet-Mulder, J. C. van (2009-2012). *COTAN Documentatie*. Amsterdam: Boom test uitgevers

Fisher, S. L., & Ford, J. K (1998). Differential effects of learner effort and goal orientation on two learning outcomes. *Personnel Psychology*, 51, 397–420.

Goldin, I., Narciss, S., Foltz, P., & Bauer, M. (2017). New directions in formative feedback in interactive learning environments. *International Journal of Artificial Intelligence in Education*, *27*(3), 385-392.

Goodman, J., Wood, R. E., & Hendrickx, M. (2004). Feedback specificity, exploration, and learning. *Journal of Applied Psychology*, 89, 248–262.

Hattie, J., & Gan, M. (2011). Instruction based on feedback. *Handbook of research on learning and instruction*, 249-271.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81 112.

Jaehnig, W., & Miller, M. L. (2007). Feedback types in programmed instruction: A systematic review. *Psychological Record*, 57, 219–232.

Kearsley, G. (2002). Is Online Learning for Everybody*? Educational Technology*, 42(1), 41-44.

Kozma, R. B. (2003). Technology and classroom practices: An international study. *Journal of research on technology in education*, 36(1), 1-14.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, *41*(4), 212218.

Law, V., & Chen, C. H. (2016). Promoting science learning in game-based learning with question prompts and feedback. *Computers & Education*, *103*, 134-143.

Leerlijnen, E. D. (2008). Over de drempels met taal en rekenen. *Hoofdrapport van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen.*

Leibold, N., & Schwarz, L. M. (2015). The Art of Giving Online Feedback. *Journal of Effective Teaching*, *15*(1), 34-46.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, *38*(1), 1-4.

Parkes, J., Abercrombie, S., & McCarty, T. (2013). Feedback sandwiches affect perception but not performance. *Advances in Health Sciences Education*, 18(3), 397-407.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. *Handbook of research on educational communications and technology*, *3*, 125-144.

Narciss, S. (2017). Conditions and effects of feedback viewed through the lens of the interactive tutoring feedback model. In *Scaling up assessment for learning in higher education* (pp. 173-189). Springer, Singapore.

Narciss, S., & Huth, K. (2006). Fostering achievement and motivation with bug-related tutoring feedback in a computer-based training for written subtraction. *Learning and Instruction*, *16*(4), 310-322.

Noordhoff Uitgevers, (2017). Retrieved from http://www.noordhoffuitgevers.nl/wps/portal/nubao/basisonderwijs2/methoden/rekenen/getalenruimtejunior

Razzaq, R., Ostrow, K. S., & Heffernan, N. T. (2020, July). Effect of Immediate Feedback on Math Achievement at the High School Level. In *International Conference on Artificial Intelligence in Education* (pp. 263-267). Springer, Cham.

Rubens, W. (2003). De (prille) geschiedenis van e-learning: omzien in verwondering. *HRD Thema*, *4*(3), 9 17.

Rubens, W., & Admiraal, W. (2003). Samenwerkend leren met behulp van ICT binnen het Europese onderwijs: ervaringen met Synergeia in vier landen. *Paper gepresenteerd tijdens de OnderwijsResearchDagen van*, *7*.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. *Psychological bulletin*, 95(3), 355.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. *Measures in health psychology: A user's portfolio. Causal and control beliefs*, *1*(1), 35-37.

Shield, G. (2000). A Critical Appraisal of Learning Technology Using Information and Communication Technologies. *Journal of technology Studies*, *26*(1), 71-79.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*(1), 153-189.

Stipek, D. J. (1993). Motivation to learn: From theory to practice.

Treffers, A., & Moor, E. D. (1990). Proeve van een nationaal programma voor het reken-wiskundeonderwijs op de basisschool. Deel 2: Basisvaardigheden en cijferen.

Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on pupils' learning outcomes: A meta-analysis. *Review of research*, *85*(4), 475-511.

Van der Kleij, F. M., Timmers, C. F., & Eggen, T. J. H. M. (2011). The effectiveness of methods for providing written feedback through a computer-based assessment for learning: A systematic review. *Cadmo*.

Vasilyeva, E., Pechenizkiy, M., & De Bra, P. (2007). Adaptation of Feedback in e-learning System at Individual and Group Level. *Proc. of PING*, 49-56.

Wang, S. L., & Wu, P. Y. (2008). The role of feedback and self-efficacy on web-based learning: The social cognitive perspective. *Computers & Education*, 51(4), 1589-1598.

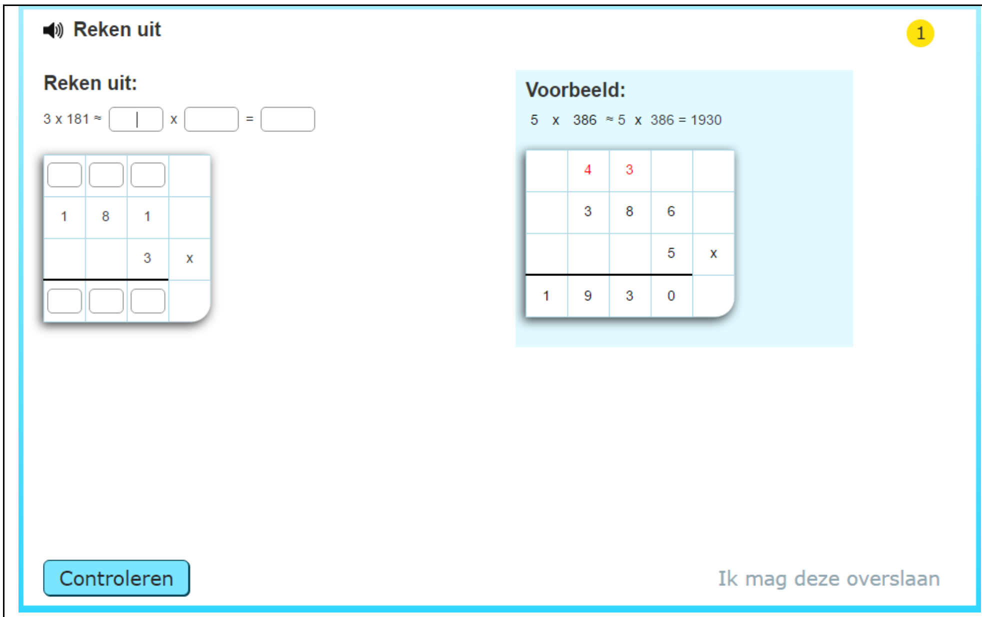
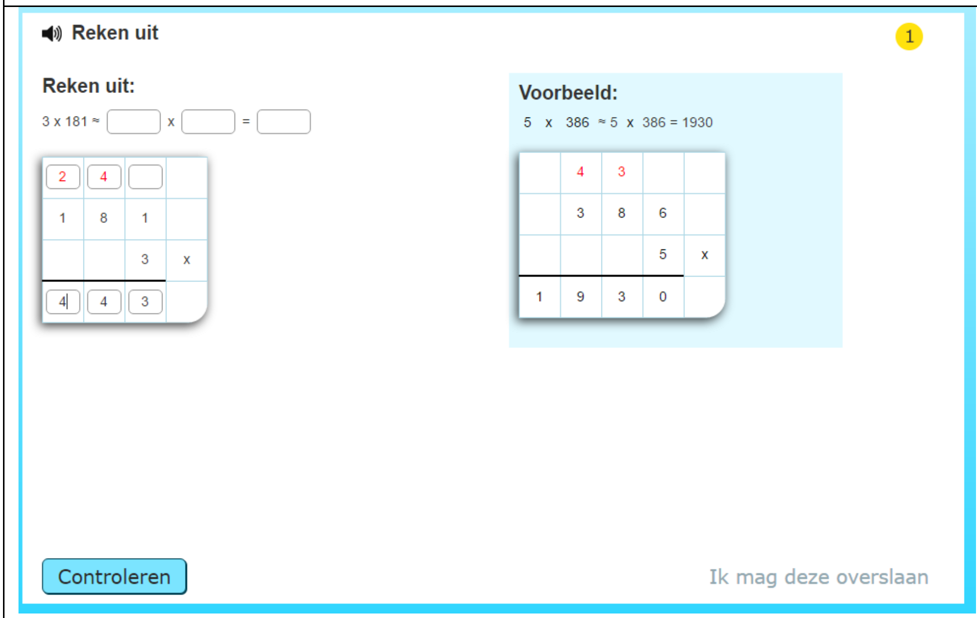
Wood, R., & Bandura, A. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. *Journal of personality and social psychology*, 56(3), 407.

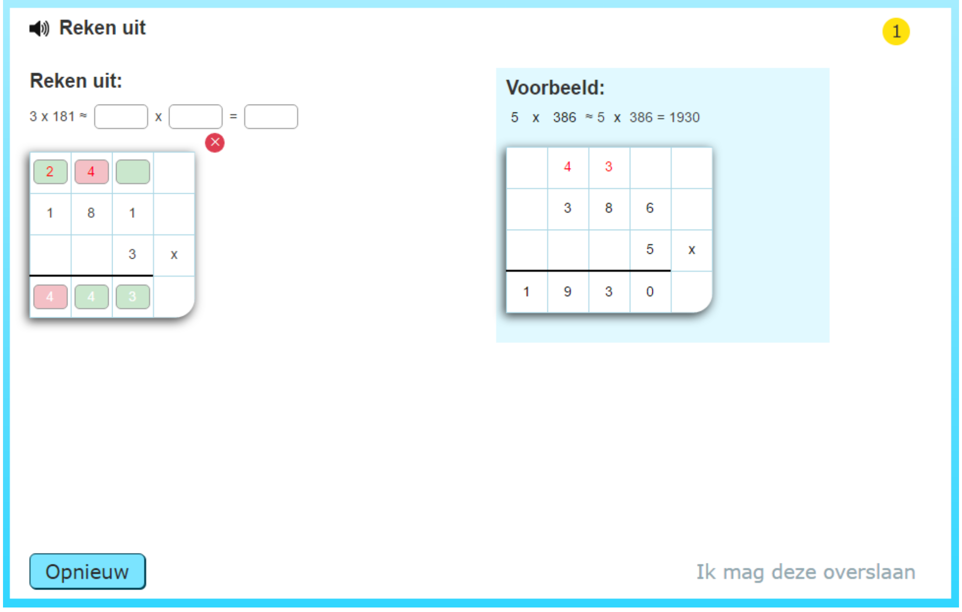
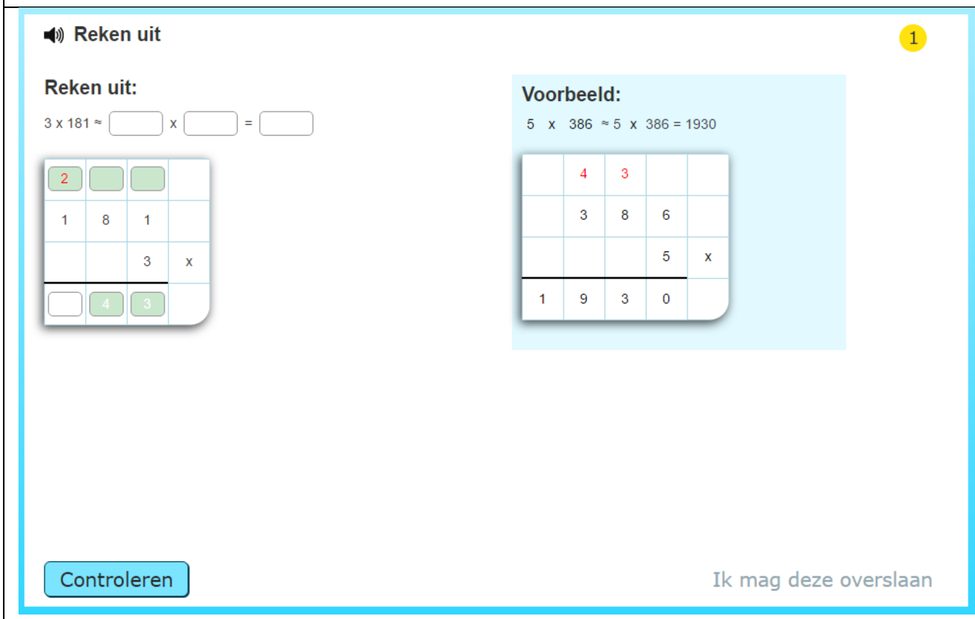
World Health Organization. (2020). Coronavirus disease 2019 (COVID-19): situation report, 72. ISO 690

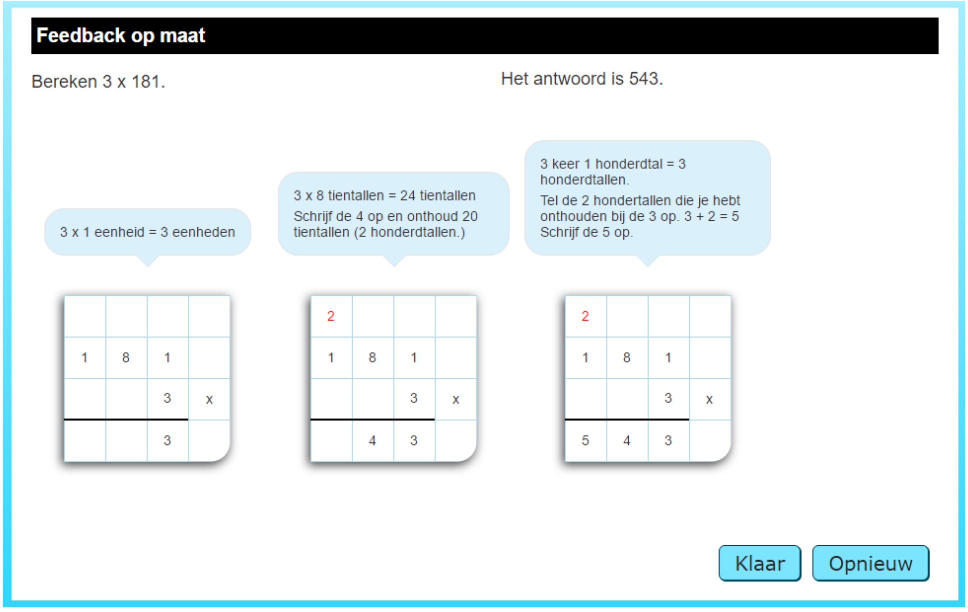
Appendices

Appendix A

Storyboard Getal & Ruimte Junior: right/wrong feedback and tailored feedback

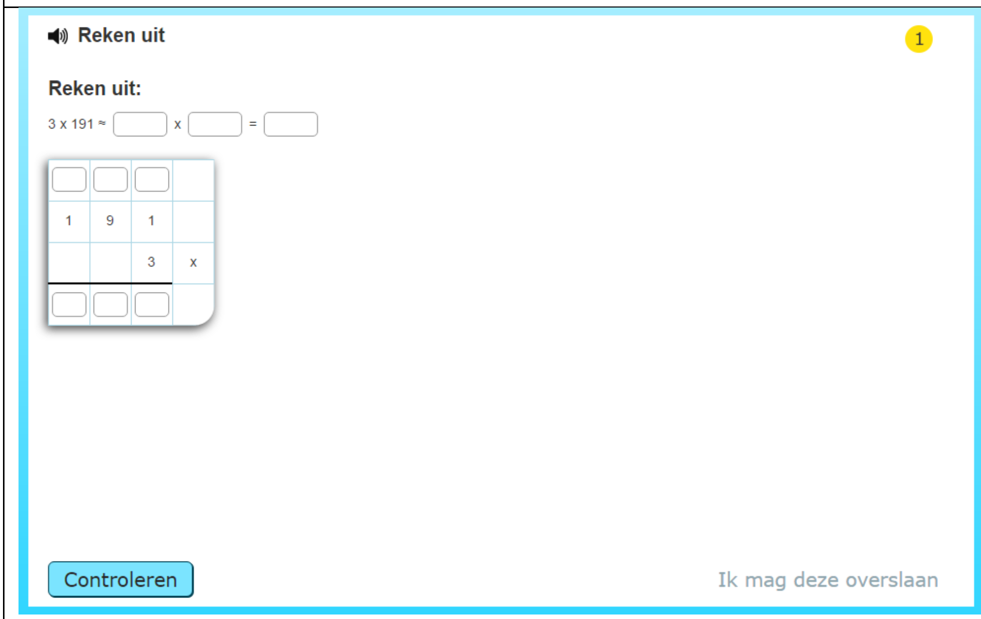
<table>
<tr>
<td>
</td>
<td>The first screen the pupil sees, when starting a new exercise.</td>
</tr>
<tr>
<td>
</td>
<td>He fills in his answers and clicks ‘check’ (controleren) to see how he has done.</td>
</tr>
</table>

	The pupil sees how well he has done the exercises by receiving right/wrong feedback. The pupil clicks ‘again’ (opnieuw) to correct his mistakes.
	After clicking ‘again’ the pupils sees this screen and is able to fill in his adjusted answers. He then clicks ‘check’ the see if he completed the exercise correctly.

<table>
<tr>
<td>
</td>
<td>Since some answers were still incorrect the pupil gets help in the form of tailored feedback.

The text is read out loud by the program.

The three images and the corresponding text appear after each other, in a timely fashion.

When the pupil is done reading the feedback he clicks ‘done’ (klaar).</td>
</tr>
<tr>
<td>
</td>
<td>After clicking ‘done’ the pupil advances to the next exercise, without being able to correct the previous exercise.

This time the exercise has no example.</td>
</tr>
</table>

Appendix B

Research invitation for the schools sent via email (in Dutch)

Beste lezer,

Voor mijn master thesis aan de Universiteit Twente, in samenwerking met Noordhoff Uitgevers, onderzoek ik het geven van feedback in een online leeromgeving. Onlangs heeft Noordhoff Uitgevers haar nieuwe rekenmethode Getal & Ruimte Junior gepresenteerd. In deze (digitale) methode kunnen leerlingen onder andere 'cijferend vermenigvuldigen' opdrachten maken en hiermee oefenen, daarbij ontvangen de leerlingen ook feedback. Naast dat leerlingen goed/fout feedback krijgen, is er in deze nieuwe methode ook ‘feedback op maat’ aanwezig. Deze nieuwe soort feedback geeft extra uitleg aan de leerling, wanneer hij of zij tweemaal achtereenvolgens een fout antwoord heeft gegeven.

Het doel van mijn onderzoek is om deze twee soorten feedback met elkaar te vergelijken en te onderzoeken welke soort feedback het meeste effect heeft op de prestaties en beleving van de leerlingen. Noordhoff Uitgevers wil graag weten of de nieuwe soort feedback, ‘feedback op maat’, beter is voor de leerlingen t.o.v. andere soorten feedback. Met uw hulp kan Noordhoff Uitgevers haar methoden verbeteren en zorgen dat leerlingen nog beter presteren.

Mijn onderzoek zal ongeveer drie kwartier tot een uur in beslag nemen. Het onderzoek zal zich richten op alle leerlingen van groep 6, die toegang hebben tot een laptop of tablet. De leerlingen worden verdeeld in twee groepen, waarvan de ene groep alleen goed/fout feedback krijgt en de andere goed/fout feedback in combinatie met ‘feedback op maat’. De leerlingen zullen eerst drie demografische vragen beantwoorden en de Tempo Toets Rekenen maken, waarna zij een digitale les cijferend vermenigvuldigen volgen. Als laatste zullen zij een vragenlijst betreffende hun beleving invullen en extra cijferend vermenigvuldigen opdrachten maken. Op deze manier kunnen wij zien of het nieuwe soort feedback effect heeft op de prestaties en beleving van de leerlingen bij het maken van deze opdrachten in een online leeromgeving.

Is het mogelijk dat ik mijn onderzoek bij u op school kan uitvoeren, dan hoor ik graag van u.

Met vriendelijke groet,

Rowan Davids
Master student aan de Universiteit van Twente
06-49740199
rowan.davids@gmail.com

Appendix C

Introduction and demographic questionnaire (in Dutch)

Voordat we beginnen met het onderzoek is het van belang dat jullie onderstaande vragen beantwoorden. Na deze vragen zullen wij beginnen met de Tempo Toets Rekenen. De regels van de Tempo Toets Rekenen zullen klassikaal worden uitgelegd. Na de Tempo Toets Rekenen zullen jullie cijfer opgaven maken in de digitale rekenmethode *Getal & Ruimte Junior*.

Vragenlijst

1. Ik ben een:

- Jongen
- Meisje

2. Ik ben geboren op (in getallen):

Dag: _____

Maand: _____

Jaar: _____

Referentienummer:

Appendix D

Tempo Toets Rekenen

Optellen

1+1 =____
2+1 =____
3+0 =____
4+1 =____
2+3 =____

7+2 =____
3+5 =____
0+7 =____
2+5 =____
4+6 =____

6+3 =____
4+3 =____
8+2 =____
3+6 =____
5+2 =____

3+8 =____
5+7 =____
2+6 =____
7+5 =____
9+4 =____

13+4 =____
7+12 =____
16+8 =____
4+15 =____
17+3 =____

6+15 =____
18+5 =____
3+14 =____
17+8 =____
7+16 =____

17+16 =____
22+13 =____
19+32 =____
34+15 =____
28+27 =____

23+28 =____
39+46 =____
65+33 =____
76+18 =____
54+27 =____

Aftrekken

2-1 =____
3-2 =____
4-2 =____
3-0 =____
5-2 =____

8-3 =____
6-0 =____
9-2 =____
7-5 =____
8-6 =____

7-4 =____
8-7 =____
7-5 =____
8-3 =____
6-5 =____

15-3 =____
13-7 =____
18-6 =____
16-9 =____
17-4 =____

18-6 =____
15-3 =____
16-8 =____
13-2 =____
19-7 =____

28-5 =____
21-9 =____
27-7 =____
25-8 =____
26-9 =____

35-17 =____
48-23 =____
26-19 =____
44-32 =____
23-18 =____

73-48 =____
54-37 =____
87-43 =____
67-49 =____
43-27 =____

Vermenigvuldigen

1x4 =____
2x2 =____
1x7 =____
0x5 =____
8x1 =____

3x10 =____
2x9 =____
4x4 =____
5x8 =____
6x0 =____

10x4 =____
3x3 =____
6x3 =____
7x3 =____
2x8 =____

6x6 =____
4x5 =____
8x4 =____
5x9 =____
7x6 =____

8x9 =____
4x7 =____
8x8 =____
7x8 =____
6x5 =____

12x4 =____
13x3 =____
7x7 =____
2x14 =____
4x16 =____

11x6 =____
7x12 =____
23x3 =____
9x9 =____
17x4 =____

4x23 =____
16x4 =____
2x36 =____
28x3 =____
5x17 =____

Delen				Gemengd	
4:2	=____	2+1	=_	2 x 1	=____
5:1	=____			2-1	=____
12:2	=____			2x5	=____
15:3	=____			4:2	=____
10:5	=____	3+2	=_	3 + 2	=____
6:3	=____			8-4	=____
20:2	=____			9:3	=____
24:3	=____			4x5	=____
36:6	=____			7+2	=____
9:3	=____			9-5	=____
24:6	=____	15:5	=_	15:5	=____
18:2	=____			3x9	=____
35:5	=____			10-3	=____
27:9	=____			5+4	=____
16:4	=____			5x5	=____
49:7	=____			8+5	=____
27:3	=____			24:4	=____
35:5	=____			13-5	=____
63:9	=____	7x4	=_	7 x 4	=____
64:8	=____	9:3	=_	9 : 3	=____
45:5	=____			17-6	=____
24:8	=____			8x6	=____
28:4	=____			6+13	=____
81:9	=____			18:3	=____
18:6	=____			19-4	=____
24:2	=____			24-6	=____
44:4	=____			15+7	=____
39:13	=____			4x13	=____
60:5	=____			33:11	=____
36:2	=____			3+19	=____
48:4	=____			36:3	=____
60:15	=____			6x14	=____
56:4	=____			43-16	=____
80:20	=____			4x16	=____
72:6	=____			37+28	=____
48:12	=____			37-29	=____
75:25	=____	42:14	=_	42 : 14	=____
52:13	=____			5x12	=____
90:30	=____	67+27	=_	67+27	=____
45:15	=____	64:32		64 : 32	=____

Appendix E

The pilot

Before the experiment, a pilot was conducted, again according to a true experimental design, i.e. one control group and one experimental group. Schools in the customer base of Noordhoff Uitgevers were contacted by phone and/or e-mail and were asked if they wanted to participate in the pilot. During the pilot ten pupils from grade three and four were evenly divided between the two groups. They had to make ten math exercises according to the same context, environment, and parameters as the experiment.

The pilot was conducted to determine which amount of math exercises the experiment should contain. However, the pilot showed different, surprising results, which did not cohere to the goal of the pilot. The results showed that the pupils from grade three failed to answer even one question correctly. On top of that, the pupils from grade four answered every question correctly in one or two tries, which means they never experienced the tailored feedback.

After talking to the pupils, it became clear that the way the exercises were presented was too difficult for pupils in grade three, but perfectly understandable for pupils in grade four. However, the difficulty of the exercises was too low for grade four. This meant that not only the difficulty, but also the way the exercises were presented determined for which grade the experiment is best suitable.

Based on the pilot was decided that the experiment would be conducted in grade four, since they experienced the way the exercises were presented in their regular curriculum. The parameters of the exercises were [110 to 450 (not divisible by 10)] x [2 to 9] = [1 to 1000] but were adjusted to [110 to 450] x [2 to 9] = [100 to 4000]. This way the difficulty was not too low, the pupils were familiar with the way the exercises were presented and, this made sure they did make mistakes and thus experienced the tailored feedback.

Appendix F

Self-efficacy questionnaire (in Dutch)

Vragen na afloop

Je hebt net cijfer opgaven gemaakt in een digitale methode. Terwijl je deze opdrachten aan het maken was kreeg je hulp van het programma, wanneer je opdrachten niet goed deed. Denk bij het maken van de volgende vragen terug aan de opdrachten en de hulp die je daarbij kreeg. Omcirkel het cijfer wat op dit moment het beste bij je past.

		volledig onjuist	**enigszins onjuist**	**enigszins juist**	**volledig juist**
1	Het lukt me altijd moeilijke problemen op te lossen, als ik er genoeg moeite voor doe.	1	2	3	4
2	Als iemand mij tegenwerkt, vind ik toch manieren om te krijgen wat ik wil.	1	2	3	4
3.	Door de hulp kon ik de opgaven goed maken.	1	2	3	4
4	Het is voor mij makkelijk om vast te houden aan mijn plannen en mijn doel te bereiken.	1	2	3	4
5	Ik vertrouw erop dat ik onverwachte gebeurtenissen doeltreffend aanpak.	1	2	3	4
6	Dankzij mijn vindingrijkheid weet ik hoe ik in onvoorziene situaties moet handelen.	1	2	3	4
7.	Ik snapte de opgaven beter door de hulp.	1	2	3	4
8	Ik kan de meeste problemen oplossen als ik er de nodige moeite voor doe.	1	2	3	4
9	Ik blijf kalm als ik voor moeilijkheden kom te staan omdat ik vertrouw op mijn vermogen om problemen op te lossen.	1	2	3	4
10.	Ik ben er zeker van dat ik de opgaven nu kan maken zonder de hulp.	1	2	3	4
11	Als ik geconfronteerd word met een probleem, heb ik meestal meerdere oplossingen.	1	2	3	4
12.	Ik kan dit soort opgaven nu goed maken.	1	2	3	4
13	Als ik in een benarde situatie zit, weet ik meestal wat ik moet doen.	1	2	3	4
14	Wat er ook gebeurt, ik kom er wel uit.	1	2	3	4
15.	Door de hulp kon ik de opgaven goed maken.	1	2	3	4

Appendix G:

Passive informed consent form (in Dutch)

Beste ouder(s), verzorger(s), voogd,

Op -dag + datum- zal op de school van uw dochter/zoon een onderzoek plaatsvinden in de klas. Dit onderzoek wordt uitgevoerd door Rowan Davids in opdracht van Noordhoff Uitgevers en de Universiteit van Twente. Onlangs heeft Noordhoff Uitgevers haar nieuwe rekenmethode Getal & Ruimte Junior gepresenteerd. In deze (digitale) methode kunnen leerlingen onder andere reken opdrachten maken en hiermee oefenen, daarbij ontvangen de leerlingen ook feedback. Naast dat leerlingen goed/fout feedback krijgen, is er in deze nieuwe methode ook 'feedback op maat' aanwezig. Deze nieuwe feedback geeft extra uitleg aan de leerling, wanneer hij of zij een fout antwoord geeft.

De bedoeling van het onderzoek is om deze twee soorten feedback te vergelijken en te onderzoeken welke soort feedback het meeste effect heeft op de prestaties van de leerlingen. Noordhoff Uitgevers wil namelijk weten of dit nieuwe 'feedback op maat' daadwerkelijk helpt bij het verhogen van de prestaties van de leerlingen. Met de hulp van uw kind kan Noordhoff Uitgevers haar methoden verbeteren en zorgen dat leerlingen nog beter presteren.

Uw kind zal tijdens het onderzoek: een demografische vragenlijst invullen, dat wil zeggen een aantal vragen beantwoorden over haar/zijn achtergrond zoals geslacht en leeftijd. Cijferend vermenigvuldigen, uw kind zal 20 cijfer opgaven maken in een digitale omgeving. Tevens zal de mening van uw kind over de gegeven feedback gevraagd worden aan de hand van een 7-punts schaal. Het onderzoek zal ongeveer drie kwartier in beslag nemen. De cijfer opgaven die uw kind zal maken zijn onderdeel van de reguliere lesstof. De verzamelde gegevens zullen volledig vertrouwelijk behandeld worden, slechts de onderzoeker heeft inzicht in de ruwe data.

Deelname aan dit onderzoek is volledig vrijwillig. Uw kind kan stoppen met het onderzoek wanneer hij/zij dit wil, zonder enige reden. Mocht u om een of andere reden niet willen dat uw kind meedoet aan het onderzoek, laat dit dan vóór -datum- weten. Dit kan door contact op te nemen met de school van uw kind. Of door contact op te nemen met de onderzoeker, zijn gegevens staan in de afsluiting. Wanneer wij niks van u horen betekent dit dat uw kind automatisch mee zal doen aan het onderzoek.

Met vriendelijke groet,

Rowan Davids
rowan.davids@gmail.com
06-49740199

Appendix H:

Post-test (in Dutch)

Controle opgaven

Je hebt net cijfer opgaven gemaakt in een digitale methode en daarna een aantal vragen beantwoord. Nu ga je ter controle onderstaande cijfer opgaven maken. Dit zijn opgaven met dezelfde moeilijkheid, maar dan zonder hulp. Je hebt bij deze opgaven dus maar één kans om het goed te doen. Probeer de opgaven zo snel en zo goed mogelijk te maken.